Ecology of *Cryptococcus neoformans* in Vietnam

Thesis

ECOLOGY OF *CRYPTOCOCCUS NEOFORMANS* IN VIETNAM

by

PHAN HAI TRIEU

A thesis submitted to the Open University U.K

For the degree of Doctor of Philosophy in the field of Life Sciences

Oxford University Clinical Research Unit

Hospital for Tropical Diseases

Ho Chi Minh City, Viet Nam

Aug, 2018

Abstract

Cryptococcal meningitis is an infection with *Cryptococcus* species; almost all disease occurs in immunosuppressed patients due to *C. neoformans*. *Cryptococcus spp* are environmental saprophytes; the ability to cause disease is considered an accident of adaptation to their ecological niche.

In Southeast Asia, cryptococcal meningitis in immunocompetent individuals is well-described and is almost always due to the *C. neoformans* ST5 (VNIa-5) lineage. I hypothesize this lineage has relative increased pathogenicity. The aims of this thesis were to: identify the ecological niche of *C. neoformans*; define the diversity of *C. neoformans* in Vietnam; compare the virulence of different genotypes, and; define transcriptional differences associated with virulence.

I identified spatial clustering of lineages VNIa-4, VNIa-5 and VNIa-32. Environmental sampling yielded 123 cryptococci the majority of which were non-pathogenic.

Using the *G. mellonella* infection model I identified an 'ecological imprint' on isolates. Within a lineage, the virulence of an isolate is determined by its source – environmental, immunosuppressed or immunocompetent patient. Moreover, the virulence of environmental *C. neoformans* can be upregulated by clinical strains of that lineage. This 'induction' effect is regulated by a protein/peptide signaling system. Comparative transcriptomics revealed 1700 genes differentially expressed between clinical and environmental isolates. Gene ontology (GO) terms relevant to metabolism and cellular biosynthesis were enriched in clinical isolates; GO terms relevant to cell cycle and cell division were significantly enriched in induced environmental isolates. Clinical and induced environmental isolates had highly similar transcriptomes. A number of genes previously confirmed as virulence determinants were upregulated in clinical and induced environmental isolates; a third of differentially expressed genes between the induced and `naive' isolates were hypothetical proteins. These are novel candidates implicated in the upregulation of

virulence. My findings provide insight into ecology of *C. neoformans* in Vietnam and the genetic control of virulence.

Acknowledgements

Studying for this PhD is really tough but rewarding. Critical thinking, but not lab experience, is the most valuable thing I have learnt from pursuing the thesis. I am grown up academically after my PhD. This project would never be possible without Prof. Jeremy Day, my principal supervisor. His demanding supervision, constructive comments and endless enthusiasm have encouraged me to finish this project. I am deeply indebted to him for this.

Next, my appreciation goes to my co-supervisor Prof. Maciej Boni for his scientific advice, insightful discussion and proofreading of earlier drafts of data chapters.

My PhD thesis would be far from completion without the help of the following individuals. I am grateful to Dr Corrine Thompson for guidance on geographical mapping and cluster analysis. I really appreciate Ms Duong Van Anh for assistance with hard lab work and Dr Phil Ashton for guidance on RNAseq. Also I acknowledge Dr Le Hoang Thanh Nhat, Dr Phung Khanh Lam and Dr Dao Nguyen Vinh for insightful discussion and guidance on biostatistical analysis. Additionally, I would like to thank Dr Lam Tuan Thanh for his scientific discussion and being my viva monitor, and Nguyen Thi Mai Trinh for phenotyping and RNA extraction. My thanks go to Dr Maia Rabaa for helpful advice and enthusiasm as my PhD monitor.

My pastime during PhD would be very tedious without amazing trips to Cat Tien, Phan Thiet and Ha Giang accompanied by my dear colleagues (Justin, Mai Thu, Van Anh, Trinh, Thanh).

Finally and most importantly, my deep gratitude goes to my parents who shape who I am today. Specially, I would like to thank my wife for her affection and giving birth to two children during my PhD.

Abbreviations

AFLP	Amplified fragment length polymorphism
AIDS	Acquired immunodeficiency syndrome
BLAST	Basic Local Alignment Search tool
bp	Base pair
cAMP	cyclic adenosine monophosphate
CD4	Cluster of Differentiation 4
CFU	Colony Forming Unit
CNS	Central nervous system
CSF	Cerebrospinal fluid
DMEM	Dulbecco's Modified Eagle's medium
DNA	Deoxyribonucleic acid
GO	Gene ontology
GXM	Glucuronoxylomannan
HCMC	Ho Chi Minh City
HIV	Human immunodeficiency virus
HTD	Hospital for Tropical Diseases
ITS	Internal transcribed spacer
L-DOPA	L-3,4-dihydroxyphenylalanine
MLST	Multilocus sequence typing
mRNA	Messenger RNA
OUCRU	Oxford University clinical research unit
PBS	Phosphate-buffered saline
PCR	Polymerase Chain Reaction
RAPD	Random amplification of polymorphic DNA
RFLP	Restriction fragment length polymorphism
RNA	Ribonucleic acid
RNAseq	RNA sequencing

RPMI	Roswell Park Memorial Institute
ST	Sequence type
tRNA	Transfer RNA
YPD	Yeast peptone dextrose
AFLP	Amplified fragment length polymorphism
AIDS	Acquired immunodeficiency syndrome
BLAST	Basic Local Alignment Search tool
bp	Base pair
cAMP	cyclic adenosine monophosphate

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 The *Cryptococcus* pathogens

The *Cryptococcus* genus comprises 70 species of pathogenic and non-pathogenic basidiomycetous yeasts. The majority of them are not harmful to humans. Of all species, *Cryptococcus neoformans* and *Cryptococcus gattii* are the most clinically relevant. *C. neoformans* causes the vast majority of cryptococcosis in HIV/AIDS patients, claiming an estimated 181,000 lives annually, with a global burden of 223,000 cases per year [1]. Infection caused by the sibling species *C. gattii* is relatively rare. This pathogen mainly infects immunocompetent patients, disease being most frequently reported from Australasia, South America and more recently the Pacific Northwest of North America, where there has been a significant outbreak of disease since 1999. *Cryptococcus* pathogens are facultative intracellular fungi in macrophages [2]. They can be found in the environment (avian excreta, soil, decaying wood) as saprophytes or in environmental micro-predators such as a free-living amoebae and nematodes) [3 4 5]

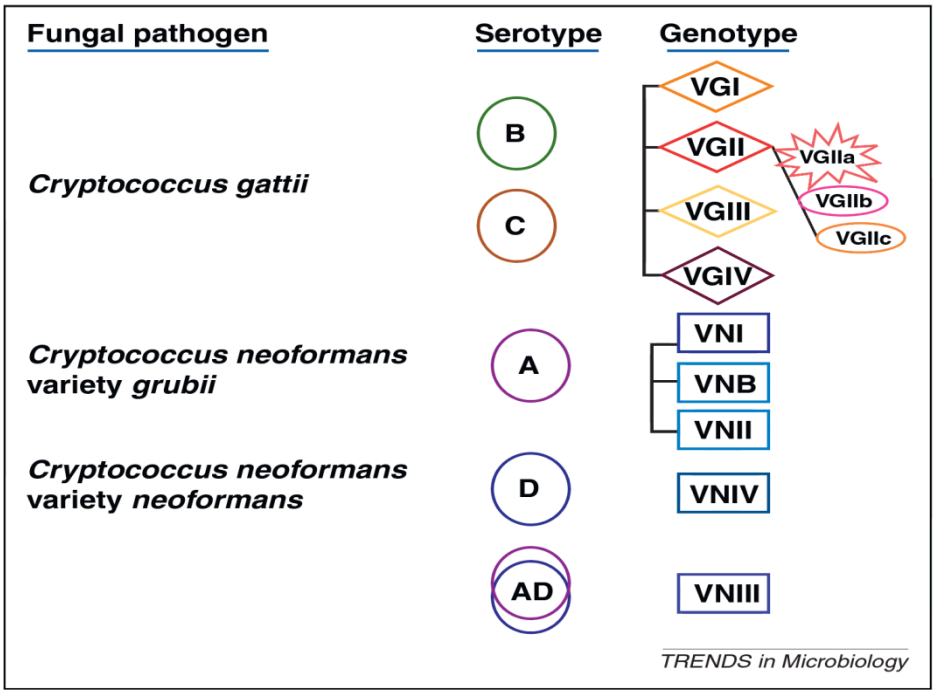


Figure 1-1 Genotypes and serotypes of *C. neoformans* and *C. gattii. Apdapted from Chaturvedi et al*[6]

Nomenclature of the (human) pathogenic *Cryptococcus* species complex is depicted in Figure 1-1. The conventional serotyping test, based on antigenic differences in major capsular glucuronoxylomannan (GXM), classifies them into 5 serotypes: *Cryptococcus neoformans* var. *neoformans* (serotype D), *Cryptococcus neoformans* var. *grubii* (serotype A), hybrid AD, and *Cryptococcus gattii* (serotypes B &C) [7] [8]. Serotyping is now infrequently performed, having been superceded by molecular methods of typing and speciation. Restriction fragment length polymorphism (RFLP) analysis of the orotidine monophosphate pyrophosphorylase (*URA5*) gene is one such reliable and easily performed method [9]. This method divides the *Cryptococcus* complex into 9 major molecular types: *C. neoformans* (VNI, VNB, VNII, VNII, VNIV) and *C. gattii* (VGI, VGII, VGIII and VGIV) [9]. The major genotypes differ in their

epidemiology, pathogenicity, disease phenotype, *in vitro* antifungal susceptibility and geographical distribution [10 7 11 12].

The most frequently used sequencing methodology is MultiLocus Sequence Typing (MLST). This is a robust sequence-based technique, which gives high discriminatory power and reproducible results. In 2007, the *Cryptococcus* Species Genotyping Working Group defined a consensus MLST scheme as the standardized method for molecular typing of *C. neoformans*, to enable description of the global population structure. Seven loci, consisting of six house-keeping genes (*GPD1, CAP59, LAC1, SOD1*, *PLB1 and URA5*) and the intergenic spacer *IGS1* were chosen and considered to have sufficient discriminatory power [13]. The MLST scheme for *C. gattii* uses slightly different loci; the major VGII molecular type has subsequently been subdivided into a, b, and c sub-types following the Pacific Northwest outbreak (Canada and Northwestern USA) using this 8 loci MLST scheme. [14]

C. neoformans sex is defined by the *MAT* locus on chromosome 4, which is over 100 kb long with more than 20 genes. The *MAT* locus encodes either the a or α allele, which determines the *MATa* or *MATα* mating type in haploid forms respectively. *C. neoformans MATα* cells are far more prevalent amongst both clinical and environmental isolates [15]. The *MATa* mating type is largely restricted to sub-Saharan Africa, and notably more frequent within the VNB population [16]. While mating was originally considered to occur between cells of opposite mating types (*MATa* and *MATα*), it is now clear that same-sex mating (*MATα* and *MATα*) can occur in *C. neoformans* var. *neoformans* [17]. During the sexual cycle, yeast cells switch to a hyphal form, but the yeast form are predominantly found in the environment and in

animal hosts. The product of mating is basidiospores. Infection is believed to occur through the inhalation of either these basidiospores, or desiccated yeasts. Basidiospores from serotype D are more resistant to environmental stresses (high temperature, oxidative stress, desiccation) and this results in them being more infectious than yeast cells [18] [19]. However, while the ability to mate is clear, it is a rare event; *C. neoformans* mostly reproduces asexually by budding, producing haploid progeny in clinical and environmental specimens [20] (Figure 1-3).

Mating can result in hybrid *C. neoformans* species. During mating, two yeast cells fuse but the two parental nuclei remain separated, forming dikaryotic hyphae (Figure 1-2). Hybridization between varieties, and even species, occurs contributing to the biological diversity of the *Cryptococcus* species complex and adaptation to fluctuating environments. The interspecies hybridization of *C. gattii* and *C. neoformans*, resulting in serotypes AB, BD and CD, has been reported from Canada and the Netherlands [21] [22] [23]. The intervarietal AD hybrid exhibits increased resistance to ultraviolet (UV) radiation and tolerance to high temperature than the haploid parents and exhibits increased pathogenicity in mice compared with serotype D [24].

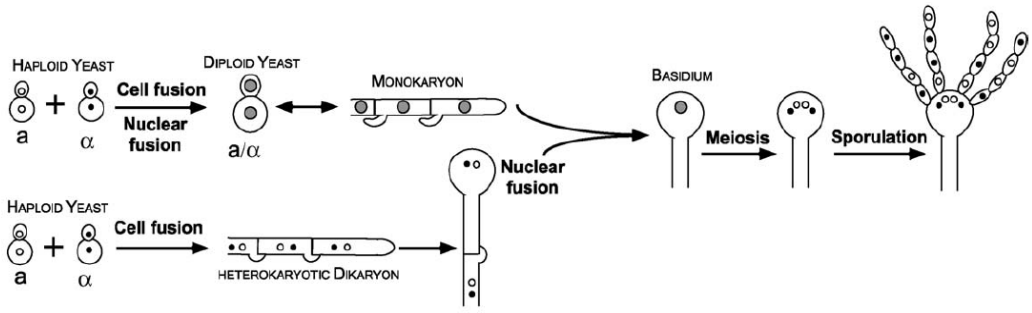


Figure 1-2 Mating of haploid yeasts causes cell fusion. Dikaryon hyphae trigger filamentous growth. Once nuclear fusion occurs, blastospores are formed through meiosis and sporulation. Adapted from Lin *et al*[23]

1.1.1 Cryptococcus neoformans

C. neoformans was first isolated from the bone of a woman suffering sarcoma in 1894, and from the environment from peach juice later in the same year. Key to the ability of the organism to cause disease is its remarkable polysaccharide capsule, a key determinant of virulence. The capsule is easily visible when Indian ink staining is applied on clinical specimen (Figure 1-3).

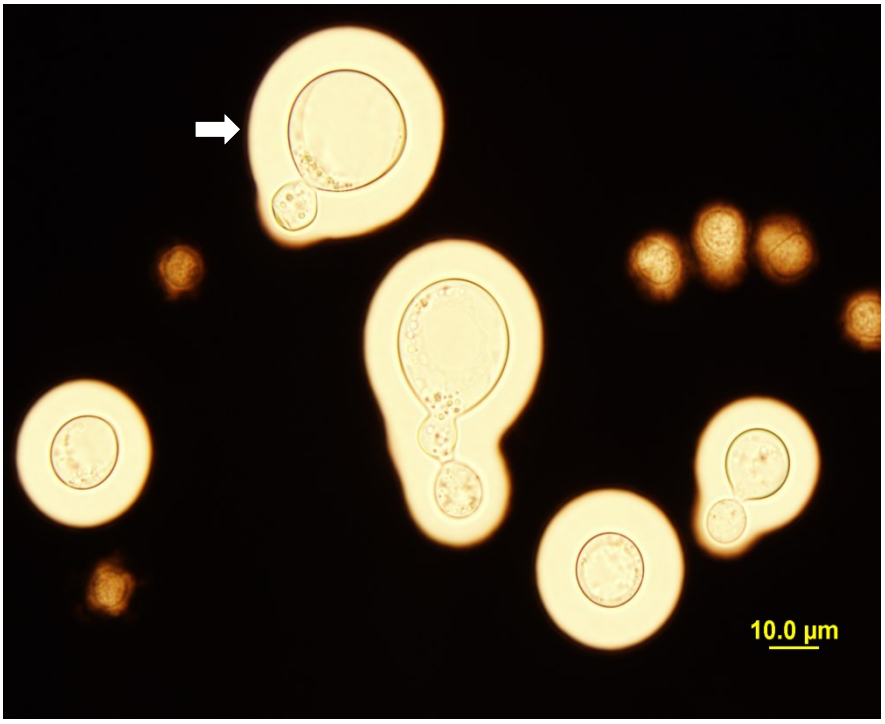


Figure 1-3 Indian ink staining of *C. neoformans* var. *grubii* isolated from *Galleria mellonella* under 1000X light microscope. Halos around spherical cells are capsule. Arrow shows a budding yeast with a daughter cell. Blurred objects in the background are hemocyte debris and body fat.

Pigeon guano is an established ecological habitat of *C. neoformans*. In the laboratory, it can grow and mate robustly on medium containing pigeon guano [25]. The pathogen has been isolated from the feet, beaks and fresh droppings of feral pigeons (*Columba livia*) [26]. However, they do not suffer cryptococcosis because their physiological

temperature is too high (average 42.5^{o}C), inhibiting fungal growth[26]. Instead, pigeons are postulated to be an important agent of dispersal for *C. neoformans* [7], acting as a mechanical vectors, particularly in densely populated urban areas.

Cryptococcal meningitis is usually seen in immunosuppressed patients and is the most common fungal infection of the central nervous system. *C. neoformans* var. *grubii* (serotype A, molecular type VNI, now renamed *C. neoformans*) is the most important strain medically, causing the vast majority (95% worldwide) of cryptococcosis in HIV/AIDS patients [7]. Driven by the HIV epidemic, VNI is the most prevalent genotype globally, except in Australia and Papua New Guinea, where *Cryptococcus gattii*, molecular group VGI prevails [27]. *Cryptococcus neoformans* var. *grubii* VNB strains are largely restricted to Botswana [28] but have been reported from Brazil and Columbia [27]. Studying a cohort of 230 HIV-associated cryptococcal meningitis cases in countries of southern Africa revealed that patients infected with lineage VNB suffered the highest mortality at one year compared with VNI and VNII [29].

Cryptococcus neoformans VNII strains can be found globally but it is rare in Europe. The hybrid VNIII, and *Cryptococcus neoformans* var. *neoformans* VNIV (serotype D, now renamed *Cryptococcus deneoformans*) are more frequently reported in Europe, (18.5 % and 18.3 % of clinical isolates, respectively)[27]. In North America, the VNII and VNIII molecular groups account for approximately 5 % of human infections [27]. The restricted distribution of VNIV to temperate regions may be attributable to the fact that VNIV strains are more susceptible to high temperature (above 40^{o}C) than VNI [30].

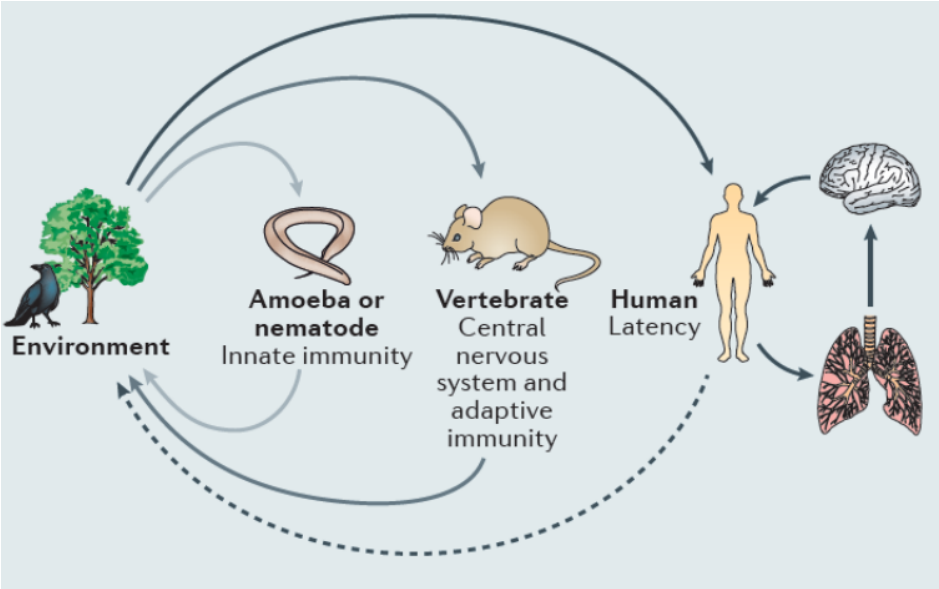


Figure 1-4 Evolution of virulence in *Cryptococcus spp. Adapted from May et al* [31]

As a saprophyte, *Cryptococcus* does not require animal hosts to complete its life cycle. Despite this, it has been found to infect a wide variety of hosts: amoebae, insects, birds, dogs, cats, dolphins, porpoises, koalas, foxes, goats, and sheep[7]. The ability to infect animals may be a 'bystander' effect that has evolved as a result of the fungal interaction with a hostile environment and the yeasts natural predators - free-living amoebae, and soil nematodes. This accidentally confers human pathogenicity (Figure 1-4). Factors that facilitate fungal survival in the environment are also required for mammalian pathogenicity. An example of such bystander dual functionality is that of the gene laccase. Laccase is involved with acquisition of nutrients from decaying food and production of melanin; melanin protects cells from solar radiation and also protects against killing by amoeba following ingestion; in

animal models it is a key virulence determinant and probably enhances intra-phagosomal survival. A further example is the capsule, which protects against predation by free-living amoebae and macrophages within the mammalian host [32 3]. Consistent with this, the the yeast transcriptional profile is similar when *C. neoformans* is exposed to phagocytosis by amoeba or murine macrophages [33]. Genes involved in non-mammalian killing (*C. elegans, Galleria mellonella*) are also required for mammalian pathogenicity [34 4 35]. Besides environmental selective pressure, microevolutionary events, such as DNA mutation, karyotype changes & chromosomal rearrangement, also contribute to driving cryptococcal pathogenesis in different hosts [36 37 38].

Human infection may be acquired by inhalation of desiccated cells or basidiospores from the environment. There is no evidence of human-to-human, or human to animal transmission. Once the fungal cells colonize the lung, they are either cleared or enter a dormant form within the lymph node complex or macrophage and evade the immune system. When human immunity is impaired, the dormant forms can be reactivated and then disseminate with propensity for the central nervous system. Cryptococcal spores are as pathogenic as yeast in the murine model although lung fungal burden is tenfold fewer in spore-infected mice compared with yeast-infected mice.[39] *C. neoformans* can cause localized infections in skin, eyes, bones, joints, lungs, etc. but the most common and severe clinical manifestation is meningoencephalitis. [20]

1.1.2 Cryptococcus gattii

In contrast to its sibling species *C. neoformans, C. gattii* mainly infects immunocompetent individuals. *C. gattii* cells, like those of *C. neoformans*, can be spherical or oval and have a polysaccharide capsule. It is associated with decaying wood or hollow of various tree species (notably *Eucalyptus calmadulensis*, but at least 50 have been associated) [6]. *C. gattii* was previously restricted to tropical and sub-tropical countries, mainly in Oceania and South America but it is now recognized that the geographical ranges have expanded notably with an outbreak of disease in otherwise healthy humans and animals on Vancouver Island in 1999. Spread to the British Columbia mainland and Pacific Northwest of the USA occurred thereafter. The incidence of disease in the outbreak on Vancouver Island rose up to 27 times higher than that seen in Northern Australia, where *C. gattii* cryptococcosis is endemic. The mortality rate has been 8.7 % [40]. In the outbreak, the majority of of disease was due to VGII, which has subsequently been divided into 3 types: VGIIa, VGIIb and VGIIc, with VGIIa causing the majority of both clinical and veterinary infections on Vancouver Island [41]. While phylogenetically related to VGIIa, VGIIc is exclusive to the Northwestern United States, particularly Oregon, and does not yet appear to have spread to Canada [14]. The enhanced virulence of the outbreak isolates correlates with an increased intracellular proliferation rate (IPR) within mammalian macrophages, and this attribute in turn is associated with significantly higher mortality rates in the mouse infection model[42]. Gene expression studies found that overrepresentation of mitochondrial transcripts correlated with higher IPR and increased virulence [42]. Interestingly, hypervirulent strains were associated with changes in mitochondria

morphology, with a tubularized rather than globular form observed in phagocytosed cryptococci[42]. The tubular morphology may be the result of mitochondrion fusion as a strategy to evade mitochondrial DNA damage as a result of oxidative bursts in the phagosome, delivering a survival benefit following phagocytosis[42].

It has been postulated that the emergence of hypervirulent VGIIa strains may be due to same-sex mating between VGIIb and a VGII isolate in Australia with spread to the United States due to the export of *Eucalyptus camaldulensis*, the major niche of *C. gattii* in Australia [43]. However, another study suggests that the global VGII population originates in Amazon rainforest, South America [44]. The original lineage of VGII then dispersed to Africa, followed by recombination event which introduced them to Australasia, Europe and North America[44].

Of the four main *Cryptococcus gattii* molecular types, VGI is the most prevalent worldwide [27]. Recently, it has been proposed that the different *C. gatti* molecular types in fact represent separate species (Table 1-1). According to this idea, the VGI molecular type is equivalent to *C. gattii* [45]. In Oceania, *C. gattii* is the most predominant pathogen with a ratio *C. gattii/C.neoformans* of 3:2 amongst clinical isolates [27]. VGIIa accounts for 39% of all cryptococcal infection in North America due to the cryptococcosis outbreak on Vancouver Island[46]. VGIII isolates are more restricted in their distribution, mainly having been recovered from Latin America (Mexico, Southern California). VGIV is rarely reported, and has mostly been isolated from HIV/AIDS patients in Africa [7]. The interspecies *C. neoformans/C. gattii* hybrid is extremely scarce, with just one isolate reported from each of Canada, Colombia, and Brazil, and three isolates from the Netherlands[27].

Current species names	PCR-fingerprinting/RFLP genotype	Proposed species names
C. neoformans* var. *grubii	**VNI, VNII**	***C. neoformans***
C. neoformans var. *neoformans*	VNIV	*C. deneoformans*
***C. neoformans* intervariety hybrid**	**VNIII**	***C. neoformans* x *C. deneoformans* hybrid**
C. gattii	VGI	*C. gattii*
	VGIII	***C. bacillisporus***
	VGII	*C. deuterogattii*
	VGIV	***C. tetragattii***
	VGIV/VGIIIc	*C. decagattii*

Table 1-1 Current and proposed species in the *C. gattii/C. neoformans* species complex. Adapted from Hagen *et al* [45]

1.1.3 Other *Cryptococcus* species

Like *C. neoformans* and *C. gattii*, other cryptococcal species are saprophytes. They are globally distributed, even having been isolated from Antarctica and the Himalayas [47]. Non-*neoformans*/*gattii* species rarely cause human infection because in general they do not tolerate the normal human physiological temperatures - temperature dependent growth is thus an important virulence factor. While capsule has been clearly demonstrated to be an important determinant of virulence of *C. neoformans*, presence of capsule per se does not necessarily confer pathogenicity. *Tremella mesenterica*, closely related to *C. neoformans,* is encapsulated but not pathogenic[48]. Another example is *C. liquefaciens*, which possesses a capsule comparable to that of *C. neoformans* in terms of size, but physical chemical analysis reveals significant differences in viscosity and elastic properties [49]. This likely explains why *C. liquefaciens* isolates are more susceptible to amoeba phagocytosis [49].

1.2 Ecology of the *C. gattii*/*C. neoformans* species complex

Cryptococci occupy a diverse range of ecological habitats, including sea, arctic, extreme pH and thermal springs, and even nucelar reactors [7]. Animal excreta and arboreal sources are established habitat of the *C. gattii/C. neoformans* species complex.

1.2.1 Trees

The ecological niche for *C. gattii* was first described when an association with the red river gum (*Eucalyptus camaldulensis)* was established in Australia[50]. This tree species has been extensively exported around the world, including to southern California, Mexico, Africa and Southeast Asia and this trade may have contributed to the global dispersal of *C. gattii*. The first environmental isolate of *C. gattii* beyond Australia was recovered from *E. camaldulensis* trees in San Francisco[51]. One historical problem with environmental sampling is that there are no reports of a systematic randomized approach to sampling, and therefore any associations are likely driven by bias in choice of both sampling location and source. Isolate collections produced through non-randomised methods cannot be considered to necessarily be representative of the true population. The discovery of the eucalyptus-associated reservoir of *C. gattii* has promoted extensive sampling of arboreal sources, and the finding, perhaps not surprisingly, that *Cryptoccocus spp* can be isolated from many tree species.

In contrast to Australia, extensive sampling of eucalyptus in India yielded very few *C. gattii*. Five *C. gattii* strains were isolated out of 315 plant debris samples in Punjab whereas none of 500 eucalyptus trees in Dehli and Uttar Pradesh gave positive result

[26]. The arboreal niche of *Cryptococcus* (serotype A and B) in northwestern India was *Syzygium cumini* (Java plum). They were repeatedly isolated in this tree species [7].

In Southern California and Southwestern USA, VGIII isolates frequently infect HIV patients. Environmental sampling identified novel arboreal hosts of this genotype: *Pinus canariensis* (Canary Island pine), *Liquidamar styraciflua* (American sweet-gum), and *Metrosideros excels* (Pohutukawa tree)[52]. In British Columbia, environmental surveillence was carried out in places where cryptococcosis outbreak had occurred. Positive soil samples were significantly higher than positive tree swabs (65% vs 35 %). High concentration of *C. gattii* in soil was associated with low moisture and low organic carbon content[53].

Environmental sampling in Colombia identified the most prevalence of serotype A, followed by B & C. Strains VGIII (serotype C) was associated with almond trees (*Terminalia catappa*). This genotype was repeatedly isolated in almond trees over 19 months. Interestingly, 96 % of VGII strains in Colombia were MATa in contrast to MATα (VGII), the most predominant isolates in Vancouver Islands outbreak[7].

Decayed wood from tree hollows in urban areas of Rio de Janeiro were colonised by *C. neoformans* var. *neoformans*[7]. Pink shower trees (*Cassia grandis*), fig trees (*Ficus microcarpa*), and November shower tree (*Senna multijuga*) remained positive one year later whereas negative trees did not become colonised[54]. This suggests perennial colonisation of *C. neoformans* var. *neoformans* in these species. A revealing study in Rio de Janeiro found contamination of domestic environment with *C. neoformans* var. *neoformans* was associated with AIDS patients with cryptococcosis. 824 samples including household dust, soil, avian droppings were

collected in 154 houses, and *C. neoformans* was detected in 20 % of the houses. AIDS patients residing in contaminated houses were more likely to develop cryptococcosis than those living in the 'clean' houses. However this result was not significant (OR=2.05, P value =0.61) [55].

A recent population genetics study has found correlation between southern Africa strains (VNI and VNB) and the indigenous African mopane tree (*Colophospermum mopane*), whereas global VNI is frequently associated with pigeon droppings[56]. Almost a third of mopane samples were positive with VNB strains. The mopane tree is important in economy and culture of southern Africa. Mopane timber is widely used in hut construction while mopane leaves are used in traditional medicine [56].Therefore, the widespread application of mopane may predispose HIV patients to cryptococcosis.

1.2.2 Animal excreta

There is a strong association between *C. neoformans* and avian excreta, suggesting that this is a primary habitat of the pathogen. *C. neoformans* was isolated in 15 % of excreta samples from macaws and parakeets in India, and 25.5 % of passerine and psittacine guano in Brazil [26] [57]. Data suggest that pigeon faeces can promote both growth and mating of *Cryptococcus spp* because cryptococcal fungi can assimilate urea and nitrogenous substances in the bird excreta [25]. However, the avian body temperature range (41.5 to 43^{o}C) is too high to support growth of *Cryptococcus spp* and therefore birds are unlikely to be internally colonized with *Cryptococcus* [7]. Pigeons and other birds are likely to be important in the dissemination of *Cryptococcus* species, with feathers, beaks, and claws being contaminated with the

yeast [26]. Avian habitat contaminated with *C. neoformans* is a plausible source of HIV/AIDS cryptococcosis in endemic areas [58] [7]. In addition to avian faeces, bat guano has been reported as a source of *C. neoformans* and *C. gattii* in India. [7]

In Thailand, cryptococcosis is the third most common opportunistic infections in AIDS patients. Environmental sampling in Chiangmai, nothern Thailand revealed *C. neoformans* was mostly associated with dove droppings samples (45 %, 45 of 100 samples)[59]. Another study detected *C. neoformans* in 24% of weathered chicken dropping samples collected in suburban areas in Phayao province, northern Thailand[57].

Environmental isolates were also detected in USA, mostly serotype A, followed by serotype D and AD. Up to 762 avian excreta samples were collected in North Carolina with 83 % of VNI [7]. In South America, bat excreta in Brazil was first reported as reservoir of both *C. neoformans* and *C. gattii*[7].

Of note, *C. neoformans* is less prevalent in Northern Russia, Canada and Scandinavia compared to tropical and temperate regions. Globally, VNII is rarely isolated in the environment[7].

1.2.3 Clinical relevance of environmental *C. neoformans*

The pathogenic potential of environmentally sourced isolates is unclear with variable findings depending on the study. According to a study in Puerto Rico, mean lethal dose (LD_{50}) of environmental *C. neoformans* is significantly higher than that of clinical isolates using intravenous murine infection model[60] . In Brazil, 69 % of environmental isolates were significantly more lethal in mice than the rest (environmental isolates)[61] whereas a recent study found that only 9% of environmental *C.*

neoformans var. *grubii* cause death 60 days post-infection in murine model[62]. Environmental isolates VGIII in Southern California exhibited increased survival time in murine model compared to MLST-matched VGIII clinical isolates[52].

Because *Cryptococcus* can cause latent infection, with periods of many years, and humans have are highly mobile with frequent travel and immigration, drawing correlations between clinical and environmental populations is complex and potentially cofounded by time and human factors. A more logical way to investigate clinical relevance of environmental isolates is to compare their genotypes with those of veterinary infections.

1.3 Cryptococcal infection

1.3.1 Cryptococcal meningitis

The most severe symptom of cryptococcal infection is cryptococcal meningitis. The clinical manifestations of cryptococcal meningitis due to infection with any species include headache, altered mental status, nausea, vomiting and fever. Visual impairment is also observed both at presentation and developing during treatment. A study from Botswana found no differences in clinical features or in-hospital mortality between *C. gattii* and *C. neoformans* infection in HIV-infected patients[63]. Sparse data suggested that *C. gattii* is more virulent than *C. neoformans* because it causes multiple lesions in lungs and brains, although disease manifestations are likely complicated by the difference in underlying host immune phenotype [64]. Mass lesions in the brain appear to be more common in *C. gattii* disease, and this may explain a higher rate of neurological sequelae in these patients [65]. Inflammatory sequelae such as hydrocephalus are reportedly more frequent with *C. gattii* than

with *C. neoformans*, although an elevated cerebrospinal fluid opening pressure is seen in 2/3 of Vietnamese patients with cryptococcal meningitis due to *C. neoformans*[66 67 64].

Cryptococcosis due to other species has been described. The most common causes of non-*neoformans*/non-*gattii* disease are *C. laurentii* and *C. albidus*, which account for approximately 80 % of cases. The remaining 20% are due to *C. adeliensis, C. humicolus, C. luteus, C. curvatus,* and *C. uniguttulatus* [7]. Non-*neoformans/gattii* pathogens mainly caused disease in patients with impaired cell-mediated immunity (immunosuppressive therapy, organ transplantation, hematologic malignancy) and patients co-infected with HIV[47]. Infection of many organ systems has been described including the skin, eyes, and gastrointestinal tract, with bloodstream and the central nervous system being the most commonly affected (39% & 32 %, respectively). The clinical outcomes are similar to those of *C. neoformans/gattii*. Amphotericin B is the mainstay of therapy because 94 % of isolates are susceptible to this drug, whereas azoles can be used for less severe or non-disseminated cases. [47]

1.3.2 Latent or acute infection

Humans are believed to become infected with *Cryptococcus* species following inhalation of spores or desiccated yeasts. Following this, the incubation period until the development of overt disease is unknown, and likely variable from patient to patient. A case of likely nosocomial transmission of *C. neoformans* has been described between two elderly patients hospitalised in beds next to each other. The isolates appeared genetically identical using PCR with four random primers and karyotyping. While both patients might simply have been coincidentally infected

with the same strain, acute transmission could not be ruled out, but currently our understanding is that person to person spread has to be considered exceptional [68]. The *C. gattii* Vancouver Island outbreak has allowed a better understanding of the incubation period for this particular strain. Disease occurred between 2 to 11 months after exposure for most patients [69]. However, there is growing evidence supporting the idea that in general most cases of disease represent reactivation of latent infection. A survey in immunocompetent children over 5 years old revealed that 70 % had serum antibodies directed against *Cryptococcus*, suggesting subclinical infection is common. Primary, self resolving pulmonary cryptococcosis could be confused with any number of pulmonary, including viral, infections in children [70]. Further evidence of latency comes from transplant patients. Here, serological studies have shown that 52 % of transplant recipients who developed cryptococcosis (*C. neoformans*) were seropositive prior to transplantation. Moreover, they developed the disease significantly earlier than patients without pre-existing cryptococcal infection [71]. In HIV infected patients, significant latency is supported by a molecular epidemiology study from France which used randomly amplified polymorphic DNA (RAPD) genotyping. Isolates from African-born patients, resident in France for a median of 10 years, had RAPD profiles that clustered together, away from isolates from patients indigenous to Europe suggesting that patients were infected with *C. neoformans* in their country of origin before migrating to France [72].

1.3.3 Single or mixed infection

Cryptococcus is ubiquitous in the environment, so humans are constantly exposed to multiple strains with different genotypes, serotypes, ploidies and mating types.

Mixed infection was reported to occur in 18 % of cryptococcosis patients infected with serotype A or D in France, however, clinical outcome did not differ significantly in these patients. These strains could have been acquired simultaneously or separately over time. Diploid strains could be co-infected environmentally or evolved from haploid strains through same-sex mating or endoreplication (duplication of nuclear DNA without mitosis) [73]. Mixed populations may enable better adaptation in the hostile host. In contrast, a study from Brazil revealed that almost all serial isolates from AIDS patients had identical PCR profiles; only a minority of patients had mixed isolates[20]. The incidence of mixed infection needs consideration when investigating the epidemiology and developing vaccines and diagnostic tools for these pathogens.

1.3.4 Cryptococcal interaction with the host immune response

After breaching the mammalian physical barrier, cryptococcal pathogens are sensed by pattern-recognition receptors (PRR) on host immune cells. PRRs are a key part of control of the innate immune system. They are are germline encoded and detect molecular motifs typical of pathogens, including fungal antigens such as β-glucan, mannan, and cell wall chitin. One examole is the C-type lectin receptors, which bind to carbohydrate ligands such as β-glucan or mannan derived from capsule and spores of *C. neoformans*[74]. Another family of PRRs is the Toll-like receptors. Toll-like receptor 4 is significant in fighting fungal pathogens, being able to recognize glucuronoxylomannan. Involvement of these receptors trigger signal transduction pathways that coordinate innate immunity, including phagocytosis and cytokine production.

1.3.4.1 Innate immunity

The complement system and phagocytic cells are crucial players in the host innate immune response to *Cryptococcus*. Complement systems opsonize cryptococcal cells in the bloodstream to attract immune effector cells and facilitate the uptake of the fungal pathogen by phagocytes. *C. neoformans* can be ingested by a variety of phagocytes. These include dendritic cells and macrophages, which following yeast ingestion, present cryptococcal antigens to T cells and produce mediators (chemokines and cytokines) to initiate adaptive immunity. Alveolar macrophages are the first line of defense against the pathogen. They are found to uptake cryptococci in murine lung[74]. However, following internalization by macrophages *C. neoformans* can survive and proliferate within acidic phagolysosomes and eventually lyse host cells. Besides lytic escape, *Cryptococcus* can exit macrophages without killing them (non-lytic exocytosis) or transfer laterally between macrophages. These mechanisms are thought to be important in dissemination of the pathogen to the central nervous system across the blood-brain barrier (the Trojan horse model)[75]. The interaction between *Cryptococcus* and macrophages may also explain how they remain latent without triggering host innate immunity.

1.3.4.2 Adaptive immunity

Whole *C. neoformans* cells or cell components (membrane, wall, proteins) trigger proliferation and maturation of naive T cells[76]. Both CD4 and CD8 T cells participate in inhibition of *C. neoformans* through direct killing or generation of cytokines that recruit and activate other phagocytes to kill invaders. CD4 T cells are crucial players of cell-mediated immunity in mice[77]. They mediate fungal clearance by recruiting

macrophages and granulocytes to murine lung. In humans CD4 T–cell deficiency (cell count < 100 cells/uL) is a predisposing factor for cryptococcosis. CD8 T cells are also involved in cryptococcal killing using granulysin. Depletion of CD8 T cells leads to reduction of survival in murine infection model[77]. Both CD4 and CD8 T cells yield Th1 cytokines, but the former is the main producer of Th2 cytokines during cryptococcal infection[76]. Th1-type CD4 T cells play important roles in adaptive immunity. They produce IL-2, IL-12, IFNγ and TNF-α. IL-12, IFNγ and TNFα provide protection against *C. neoformans* in mice. Human studies confirm protective roles of IL-2, IL-12 and IFNγ against cryptococcal infection[76]. The presence of the IFNγ and/orTNFα CD4 T-cell response is associated with survival of HIV patients with cryptococcal meningitis. Prescription of IFN-γ to the standard treatment regimen increases the rate of fungal clearance from CSF[77] in HIV infected patients. In contrast, higher levels of Th2-associated cytokines such as IL-5 and IL-13 are associated with worse outomes from cryptococcosis[78]. They are associated with increased lung burdens and murine susceptibility to *C. neoformans*[76]. Absence of an IL-13 response correlated with higher yields of protective IL-17 and IFNγ cytokines. B-cells also play crucial roles in protection against cryptococcosis - deficiency of B cells and antibody increases the risk of cryptococcal infection in both HIV infected and uninfected patients[76]. This has been elegantly demonstrated in the mouse model where in mice with severe combined immunodeficiency T-cell transfusions from B-cell-depleted mice did not result in effective adaptive immunity to *C. neoformans* in the recipients [77].

1.3.5 Epidemiology of cryptococcosis

1.3.5.1 HIV-associated cryptococcosis

Globally cryptococcal meningitis (CM) is responsible for 15 % of HIV-associated deaths [1]. The vast majority of these cases was caused by *C. neoformans* var. *grubii* (serotype A). *C. gattii* meningitis in AIDS patients is still rare, mostly caused by VGIII and VGIV, accounting for 2.4% – 30 %, in South America and Southern Africa [79]. The greatest burden of CM is in sub-Saharan Africa (73 % of total) with annual prevalence of an estimated 162,500 cases. The region suffering the second greatest incidence is Asia and the Pacific (19% of total) with an annual incidence of 43,200 cases. Clinical outcomes remain unacceptably poor. The mortality rate by 3 months can reach 50 to 70 % in sub-Saharan Africa, and 15-26 % in USA and France for 10 weeks[64] . Differences in mortality are in part due to the availability of antifungal and antiretroviral therapy, although incidence and mortality rates have not fallen significantly in South Africa over recent years despite improvements in HIV treatment access [1]. In addition, mortality rates may exacerbated by cryptococcal immune reconstitution inflammatory syndrome (IRIS), occurring in 6-30 % patients following initiation of ARVs (antiretrovirals) in HIV co-infection. In IRIS, antiretroviral therapy restores cryptococcus-specific immune response; this, in turn, elicits paradoxical clinical worsening of preexisting infection [80].

As an opportunistic infection, *C. neoformans* infection is driven by the HIV pandemic. When AIDS expanded since 1980s, CM emerged and infected 5-10 % AIDS patients. In France, incidence of HIV-associated CM increased fivefold from 1985 to 1993 while the number of cases in HIV-negative individuals remained stable[79]. However,

the upsurge of cryptococcosis has receded in developed countries (France, UK, USA) since the adoption of antiretroviral therapy (ART) and uses of azoles in 1990s [80]. The incidence of HIV-associated CM in France dropped by 46 % between 1997-2001 period. Incidence varies with geographical regions probably due to exposure in the environment. For example, cryptococcal prevalence is higher in northeastern Thailand (24 %) than southern Thailand (7%) (P value < 0.001) [81].

The prime risk factor for HIV-associated cryptococcosis is profoundly suppressed immunity (CD4 cell count < 100 cells/μL). Predictors for mortality at 2 weeks after beginning treatment include: abnormal mental status, high fungal burden in CSF, and high peripheral white blood cell count, older age (>50) and fluconazole-based induction treatment, suggesting fungicidal amphotericin-based and earlier diagnosis are important to improving outcomes [82].

1.3.5.2 Non-HIV cryptococcosis

Besides HIV patients, those who have other causes of defective T cell-mediated immunity are also at increased risk of cryptococcosis. Such conditions include: malignancy, solid organ transplant recipients and autoimmudiseases. Cryptococcosis is the third most common fungal infection in solid-organ transplant in the USA (2004-2007), accounting for 2.8 – 8 % of transplants [83] [79]. The etiology is considered to be most likely reactivation of dormant infection. Expansion in the use of immunosuppressant drugs (tacrolimus, cyclosoporin, calcineurin inhibitors, cytoxic agents) to prevent graft rejection and cancer treatment contribute to the increased incidence of cryptococcosis, especially in developed countries, with the median time to diagnosis being 20 months after transplantation [84]. In the USA, kidney transplant

receivers suffer the highest incidence of cryptococcosis in a retrospective study from 1996 to 2010.[85]

Other groups of non-HIV, nontransplant recipients with cryptococcosis include those prescribed with immunosuppressive agents & patients with hematological diseases[84]. *C. gattii* predominantly infects immunocompetent hosts. Molecular epidemiology varies with geographical regions. In Australia and Papua New Ginea VGI is the most prevalent with mortality whereas in North America (British Columbia, Oregon and Washington state), VGII (VGIIa, VGIIb and VGIIc) is exclusivel detected.

1.3.6 Epidemiology of Cryptococcosis in Vietnam

Globally *C. neoformans* var. *grubii* VNI causes the vast majority (>90%) of cryptococcosis in AIDS/HIV patients. This is true in Vietnam, where it is responsible for 100% of disease in HIV infected patients at the Hospital for Tropical Diseases, Ho Chi Minh City. However, in Vietnam, this strain is also responsible for 73 % of infection in non-HIV infected patients [86]. The majority of these patients have no clear underlying immunosuppressive disease. Using amplified fragment length polymorphism (AFLP), Vietnamese *C. neoformans* var. *grubii* strains were further separated into two distinct lineages: VNIϒ and VNIδ (Figure 1-6). VNIϒ strains, subsequently identified as multi-locus sequence type (ST5) caused 84 % of *C. neoformans* infection in HIV-uninfected patients, but only 38% of infection in HIV-infection patients (odds ratio, 8.30; 95% CI, 3.04 to 26.6; P < 0.0001). Moreover, in HIV-uninfected patients, underlying diseases were more common in those infected with VNIδ (odds ratio, 6.43; 95% CI, 0.75 to 85.1; P = 0.05). Therefore, it is

hypothesized that VNIɣ has increased pathogenicity (ability to cause disease) compared to other VNI clades. This finding is consistent with those from China, the neighbour to the north of Vietnam. A study from 16 mainland provinces in predominantly HIV uninfected patients, found that most infections were due to a distinct cluster within the VNI clade [87].

Since *C. neoformans* is primarily an environmental saprophyte, this increased pathogenicity presumably represents an adaptation to a particular ecological niche which coincidentally confers increased pathogenicity. An alternative hypothesis would be that these apparently immunocompetent patients may have a subtle and highly *Cryptococcus*-specific immunodeficiency which predisposes them to cryptococcosis.

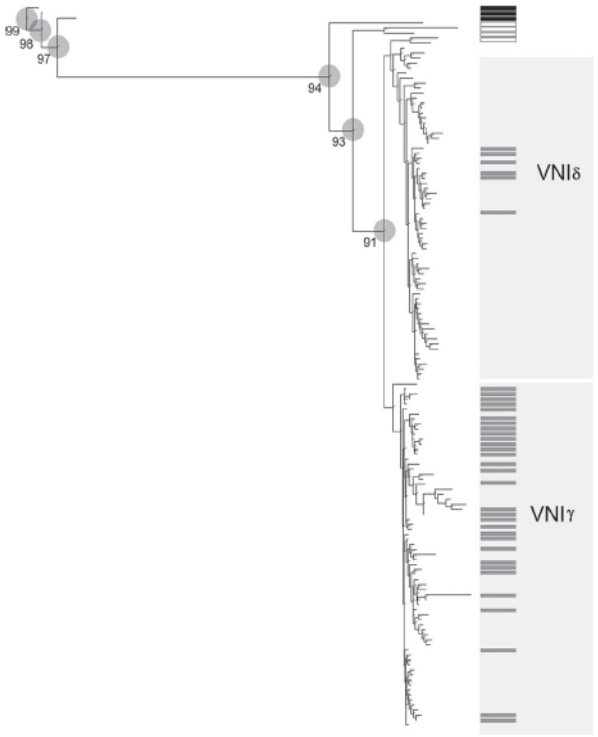


Figure 1-5 AFLP derived neighbour joining tree for Vietnamese VNI isolates. Gray bars, isolates from HIV-negative patients; white bars, *C. neoformans* control; black bars*, C. gattii* control. Adapted from Day et al [86]

Comparison of clinical phenotypes by lineage in HIV patients demonstrated that patients infected with ST5 had significantly lower baseline fungal burden in CSF (P=0.01) but higher blood lymphocyte cell count (P=0.001) than those infected with non-ST5. ST5 strain infections were associated with worse disability outcomes for survivors at 70 days than non-ST5 did (P=0.046) [88]

1.3.7 Antifungal treatment

Cryptococcal meningitis, if untreated, is invariably fatal. The standard therapy includes three phases: induction, consolidation and maintenance. Amphotericin B (AMB), a fungicidal antifungal, should be administered intravenously in 14 days for induction phase, with a dose of 0.7 – 1 mg/kg. AmB can be replaced with liposomal AmB (3-6 mg/kg) to minimize renal side effects.

To enhance fungicidal efficacy, AMB (1 mg/kg/day) should be prescribed with Flucytosine (5-FC) at a dose of 100 mg/kg/day for two weeks. This regimen reduces mortality at 70 days compared to AMB monotherapy and significantly increases rate of fungal clearance in CSF.[67] Omission of 5-FC has been associated with late relapse or treatment failure [79]. Subsequently, fluconazole, a fungistatic agent, should be used in consolidation phase (400 mg/kg/day) for further 8 weeks. Consolidation phase is followed by maintenance regimen with fluconazole (200 mg/kg/day). Overall, the gold regimen is AmB (1mg/kg) + 5-FC (100mg/kg) for two-week induction, followed by fluconazole (400 mg/day) for 8 weeks and maintained by fluconazole (200 mg/day) until HIV is controlled by ART [79] [67].

There is little high quality data to guide the treatment of *C. gattii* infection. Significant differences in susceptibility of serotypes was not detected between *Cryptococcus spp* in a study including 128 isolates (86 *C. neoformans* and 42 *C. gattii*) [89]. Therefore, the Infectious Disease Society of America and WHO recommended the same treatment regimen for these pathogens. Because the risk of development of cryptococcomas is considered increased in disease due to *C. gattii*, it is recommended that duration of treatment is prolonged [79]. Targets of antifungals are illustrated in Figure 1-6.

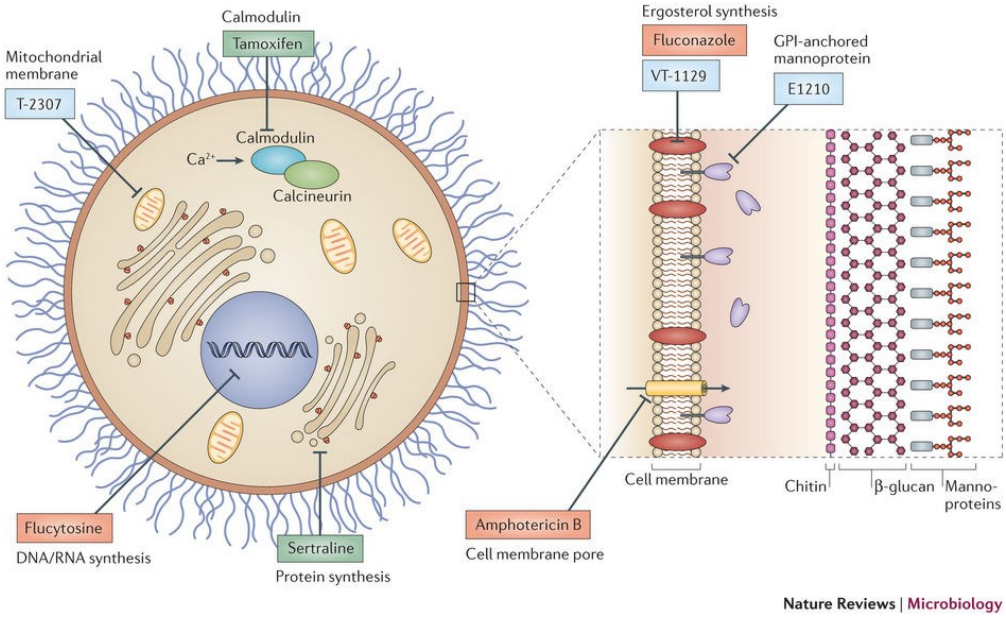


Figure 1-6 Schematic image of cryptococcal cells and drug targets. AMB generates cell membrane pores by binding to ergosterol of cell wall. 5-FC inhibits DNA and protein synthesis. Fluconazole interferes with ergosterol synthesis, weakening cell membrane. B-glucan, chitin and mannoprotein are component of capsule. Taken from Kronstad *et al* [90]

1.4 Virulence factors

Cryptococcosis occurs as a result of interaction between fungal virulence factors and susceptible hosts. Cryptococci do not rely on a single virulence factor for pathogenicity but rather require a minimum set of interacting properties for pathogenicity. These virulence factors necessary to cause human disease are complex phenotypes, selected by environmental pressures likely driven by the need to evade predators in the normal ecological niche such as nematodes and amoebae [91]. Understanding virulence factors is not only interesting in elucidating pathogenesis; virulence factors are themselves potential drug targets.

1.4.1 Thermotolerance

Susceptibility to human physiological temperature is a prerequisite for any invasive pathogen. Only a very few of 1.5 million fungal species are thermotolerant (35-40^{o}C) [92]. While there are phylogenetic relatives of *C. neoformans* which possess virulence determinants similar to those seen in *Cryptococcus*, such as capsule and melanin production, they are non-pathogen in humans simply because they cannot tolerate human body temperature [93]. Such an example is *C. podzolicus*. Within *Cryptococcus*, there is variability within the expression of particular virulence factors. For example, *C. neoformans* var. *neoformans* (serotype D) is more susceptible to elevated environmental temperature than other species [30]. This may explain its restricted distribution in northern Europe. The genes associated with thermotolerance are potential drug targets. A number of genes involved in signal transduction pathways are necessary for growth at 37 ^{O}C including *RAS1, MPK1* and *CTS1* and Calcineurin A (*CNA1*). The *CNA1* site-directed mutant is not pathogenic in the rabbit model [94].

Different virulence factors can be closely linked. For example, the genes *VHP1* and *UGD1* are essential for the synthesis of capsular components and also important for thermotolerance. Other genes, such as *STE20* and *CPA1*, are required for growth at elevated temperature (35-40^{o}C) [92].

1.4.2 Nutrient acquisition

Following phagocytosis by macrophages or free-living amoebae, the yeast cell enters a highly hostile environment, being, deprived of nutrients (carbon and amino acids), and facing oxidative stress and antimicrobial proteins [33]. The phagosome is then acidified and fused with lysosomes, the purpose being to destroy the invaders. To adapt to nutrient scarcity, *C. neoformans* increases uptake of host nutrients by expressing multiple membrane transporter genes for sugars, nitrogen and amino acid (*DIP5, ITR1, AMT1, AMT2*). This adaptation is seen in both amoebae and the macrophage model [33]. Genes encoding for key enzymes of the glyoxylate cycle (isocitrate lyase, malate synthase) are also upregulated to provide substrates for gluconeogenesis in place of glycolysis. Iron transporter and iron permease expression are also induced in internalized cells [95]. Overall, approximately 14 % of genes regulated in response to phagocytosis by murine macrophages were associated with nutrient transporters [33]. Furthermore, *C. neoformans* can resort to autophagy, a mechanism of recycling defective organelles and cellular components, to circumvent nutrient starvation in the phagosome. The autophagy regulating genes (*ATG3*, *ATG9*) are expressed when cryptococci are ingested by murine macrophages [95]. *ATG8* mutants exhibited attenuated virulence in a mouse model of infection [96].

1.4.3 Capsule

As might be expected, the capsule of *Cryptococcus* is an important virulence factor. Acapsular mutants (*CAP59* knockout gene) have attenuated virulence in the murine model [97] and acapsular cells exhibit reduced survival compared with encapsulated cells in both amoeba and the *Galleria mellonella* infection model [3] [98]. In the environment, capsule is postulated to have numerous functions. First, it is believed to protect cells from desiccation. However, often capsule is poorly produced in the environment; the resultant smaller diameter may confer an advantage in aerosolistaion and improve the efficacy of infection through inhalation by hosts.

During intracellular parasitism, capsule enlargement is induced to confer protection against reactive oxygen species from H_2O_2, antimicrobial peptides (e.g. defensin) and Amphotericin B [99]. Encapsulated cells were less sensitive to AMB (0.06 - 0.25 ug/mL) than smaller cells *in vitro*[99]. Capsular expression (in term of size) varies between strains, organs infected, and even individual isolates. The lung is particularly associated with highly encapsulated cells [100]. Encapsulated cells are also found in CSF and in phagocytes (macrophages, amoebae, larval hemocytes) but capsule is almost invisible under microscope and is best revealed using counter-staining. When visualised by such staining with Indian ink, the capsule is identified as a clear halo around the individual yeast cells (Fig. 1-3). Experimentally, capsule production can be promoted by culturing cells in minimal media: DMEM (Dulbecco's Modified Eagle's Medium) with rich CO_2 gas (5%) and bovine fetal serum.

Capsule composition is heterogenous, consisting of polysaccharide glucuronoxylomannan (GXM, 90-95 % of capsular mass),

glucuronoxylomannogalactan (GalXM, 5-8%) and a small proportion of mannoproteins (MP,) [100]. GalXM is of particular interest because of its role in the immune response. Additional polysaccharides are chitin and β-glucan (Figure 1-5). When mannosylated, mannoproteins are important cryptococcal antigens implicated in eliciting the immune response in mice. The capsule has an antiphagocytic function. Encapsulated cryptococci evade phagocytosis in the absence of opsonization and subvert phagosome digestion. Capsular enlargement itself acts as a physical barrier to phagocytic ingestion. A thick capsule enhances intracellular survival although a thin capsule is thought to promote better traversal of the blood-brain barrier in *C. gattii* [101]. When internalized in the phagosome, cryptococci shed GXM polysaccharide into vesicles, filling the cytoplasm of the infected macrophage. This combines with the leaky phagosomal membrane and eventually leads to rupture of the macrophage by accumulation of replicated cryptococci [102]. Capsule can quench reactive oxygen species and other antimicrobials. The first discovered capsular gene was *CAP59*, encoding transmembrane protein involved in GXM export. Another capsular-associated gene is *CAP64*. Deletion of *CAP64* gene leads to avirulent strains. Capsular expression of acapsular serotype A strains can be rescued by complementing *CAP64* from serotype D [15]. Other genes contributing to capsular formation include *CAP60* and *CAP10*, their importance for virulence having been demonstrated with knock-out/complementation studies [15] [103]. Further genes important for acetylation and xylosylation of GXM include: *CAS31, CAS32, CAS33, CAS34, CAS35, UGD1 and UXS1* [15].

Capsule polysaccharide may be important in the pathophysiology of raised intracranial pressure, an important complication of cryptococcal meningitis that can result in visual impairment and death. This is likely mediated through shed capsule resulting in increased CSF osmolarity, although inflammation likely contributes [66].

1.4.4 Melanin

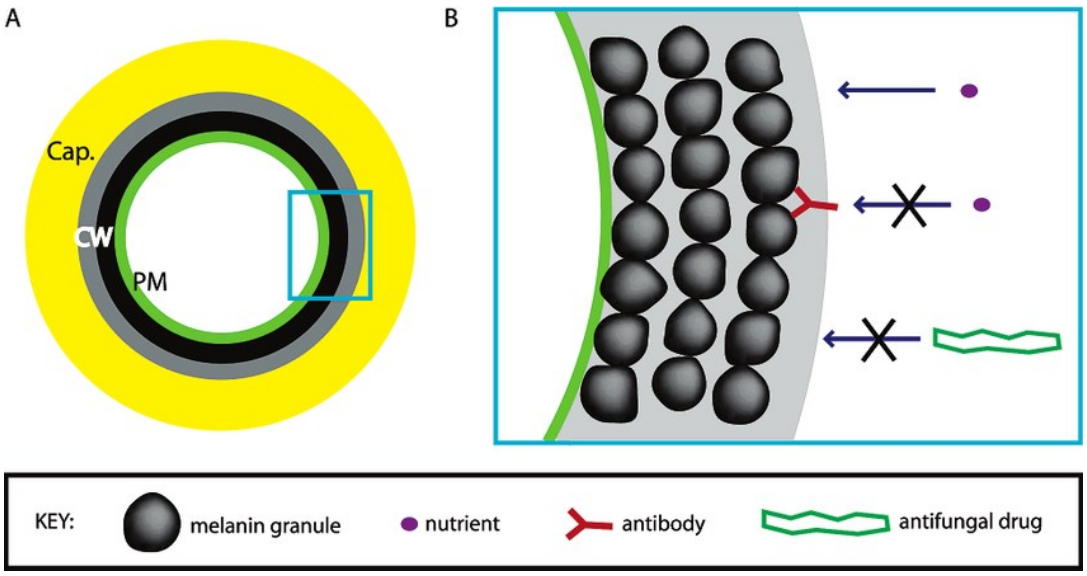


Figure 1-7 Schematic representation of melanin. (A) cross-section depiction of *C. neoformans.* Yellow line is capsule (Cap) surrounding the cell. The next layer is cell wall (CW). Melanin is embedded in CW and next to plasma membrane. (B) Menalin particles form lay layers. They allow passage of nutrients (sugar and amino acid) but hinder entry of antifungal aggregate (AMB). Antibody can bind to melanin and block entry of nutrients. Adapted from Eisenman *et al*[104]

Melanin is a virulence factor for many microbial pathogens: *C. neoformans, Aspergillus spp., Burkholderia cepacia* [105]. Melanin, localised in the cell wall, is produced from phenolic compounds (Figure 1-7). Laccase genes, encoded by *LAC1* and *LAC2*, are key in melanin production. In the environment, melanin probably confers resistance to UV radiation, high temperatures, heavy metals and protects against predators (nematodes and amoebae)[104]. Melanized cells are more resistant

to oxidants[15] and killing by antifungal drugs (caspofungin and AMB) *in vitro* [106]. Melanin-deficient *LAC1* mutant strains are less virulent in mice than wild type strains [107]. Together, melanin poses challenges in the antifungal treatment of cryptococcosis with AMB.

Most non-*neoformans* cryptococci lack laccase activity (Figure 7) and are not pathogenic. An exception is *C. podzolicus*, closely related to *C. gattii*, which exhibits strong laccase activity and produces capsule, but is avirulent probaly because of high susceptibility to temperatures of 37 ^{o}C [93]. Melanization is also variable between strains. A survey of *C. neoformans* (VNB genotype) derived from sub-Saharan Africa showed that VNBI isolates are more heavily melanized than VNBII, and environmental VNBI are significantly more melanized than clinical counterparts [16].

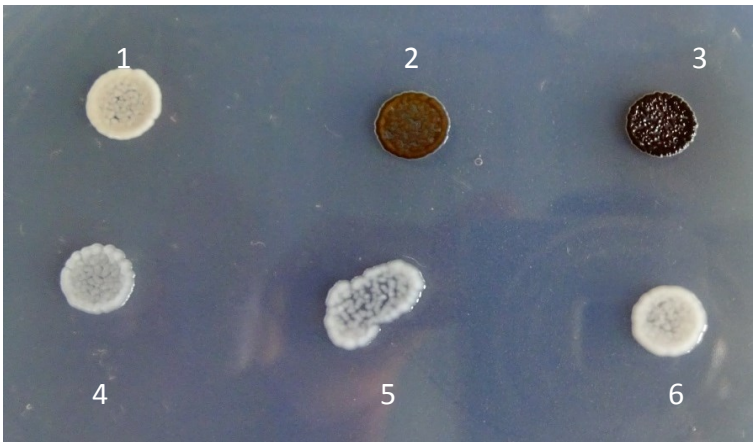


Figure 1-8 Melanin production of *Cryptococcus spp* on L-DOPA agar, melanin-inducing media. 2 &3, environmental *C. neoformans*; 1 C. *laurentii*; 4&5, C. *hevanensis*; 6, C. *rajasthanensis*

1.4.5 Oxidative defense

Following internalization, cryptococci are exposed to attack by oxygen radicals, nitrogen oxide and antimicrobial proteins. Hence oxidative defense is essential for fungal survival in phagocytes. Nitrous Oxide (NO) has a fungistatic effect against *C.*

neoformans in vitro. Cryptococcal NO-detoxifying enzymes include flavohemoglobin denitrosylase and S-nitrosoglutathione (GSNO) reductase. These enzymes are encoded by *FHB1* and *GNO1*. *FHB1* mutants demonstrated attenuated virulence in the mouse model, whereas *GNO1*-deficient strains did not. However, the *FHB1 GNO1* double mutant was less virulent than single mutants of these genes [108], suggesting *GNO1* plays an ancillary role in counteracting nitrosative stress. As might be expected, the genes involved in oxidative and nitrosative stress are induced following phagocytosis in amoeba and macrophages. Catalase and cytochrome c peroxidase take part in detoxification of H_2O_2. However, four catalase isoforms are dispensable for virulence in murine model, suggesting there is considerable redundancy of the antioxidant systems in *C. neoformans*. The thioredoxin reductase enzyme (Trr1) associated with resistance to nitrosative stress is also upregulated upon phagocytosis [33]. Other enzymes involved in counteracting the oxidative burst are two isoforms of superoxide dismutases (SOD), glutathione peroxidases, the inositol phospho- sphingolipid-phospholipase C1 (Isc1) and the protein kinase C (Pkc1) [109]

1.4.6 Other virulence factors

A number of other enzymes and proteins that facilitate cryptococcal survival under environmental stresses have been identified. One of the most common stresses is alteration of extracellular pH. *C. neoformans* grows well at human physiological pH (pH 7.4) and adapts to the acidic phagolysosome (pH 5), but suffer severe growth impairment at alkaline pH of 8 or above. The transcriptional factor RIM101 is involved in regulating cryptococcal pH response. RIM101-deficient strains are more

susceptible to the presence of alkaline pH, elevated salt concentration, and an iron-deprived environment [110]. RIM101 mutants also exhibit remarkably defective capsule growth at the cell surface by secreting capsular materials rather than maintaining them[111] .Paradoxically, unlike other capsule-deficient strains, hypocapsular RIM101 strain is hypervirulent compared with wild type H99 in the murine inhalation model. Intracellular survival of RIM101 was also better than H99 within macrophages[110]. Another study found the RIM101 mutant strain maintained virulence in murine model compared with H99. This could be explained by the fact that RIM101 mutant sheds four times as much GXM into the medium as wild type [111]. The increased level of capsular GXM causes immunosuppressive effect.

Another virulence–associated protein called DEAD-box is encoded by *VAD1* gene. *VAD1* mutants are rapidly cleared from lung and fail to establish dissemination to the brain and spleen in mice. The gene facilitates cryptococcal evasion of the host innate immune response [112].

The *APP1* gene mediates antiphagocytic activity independent of capsule. The product is excreted extracellularly and can be found in patient serum. APP1 protein binds to complement receptors 2 and 3, thereby arresting complement-mediated opsonization and inhibiting phagocytosis [113] [114]. Another capsule-independent antiphagocytic protein in the GATA family of zinc-finger transcriptional factor is encoded by *GAT201*. The virulence of the *GAT201* knockout mutant is attenuated in the murine model compared to wild type. The capsule of *GAT201* mutants is not visible using India ink counterstaining and can be detected by GXM-specific monoclonal antibody only. Moreover, *GAT201*-deficient strains are phagocytosed

more profoundly than *CAP* gene mutant strains (*CAP10, CAP60* and *CAP64*). Double mutants (*CAPΔ* & *GAT201Δ*) are internalized more dramatically than single CAP mutants, implying the defective evasion of phagocytosis by *GAT201* mutants is caused by both capsule-dependant and capsule-independant mechanisms [115]. *GAT201* is a key virulence regulator and controls the induction of 1100 genes (16 % of the genome) with 62 genes having promoters bound to *GAT201* for gene expression in tissue culture conditions [116]

1.4.6.1 Secreted vesicles

GXM, a major component of capsule, is produced intracellularly and transported to extracellular space for capsular assembly via secretory vesicles [117]. GXM is a potent immunosuppressant, mediating apoptosis of T cells via overexpression of FasL death receptor in antigen presenting cells [118]. Apart form GXM, *C. neoformans* also secretes a heterogenous group of extracellular vesicles. Using proteomic and biochemical analyses, other virulence determinants were detected in the vesicles: laccase, urease, phosphatase, enzymes related to capsular production, antioxidant proteins (superoxide dismutase, thioredoxin reductase, catalase A) [119]. Secretory vesicles are produced by *C. neoformans* during during macrophage infection, suggesting they have a key role in cryptococcal pathogenesis.

1.4.6.2 Morphological switches

Morphogenesis contributes to pathogenicity. A cryptococcal strategy to evade immune recognition is to change the surface presented to host immune systems. Morphological switches differ from spontaneous mutations, which also lead to divergent phenotypes, in that they are reversible and readily identified in a fraction

of the population [120]. There are three types of morphological changes: phenotypic switching, pseudohyphae and titan cells. Phenotypic switching was first described in a strain of *C. neoformans* displaying three colony types (serrated, wrinkled and smooth) with altered virulence in the murine model [121]. Mice infected with wrinkled colonies suffer the highest mortality. The difference in colony types and virulence was associated with minor karyotype changes [120]. Another study found phenotypic switching in mice, with smooth and mucoid colonies deriving from a strain. Mucoid colonies exhibited larger GXM, reduced phagocytosis compared with smooth cells and consequently increased virulence in mice [122]. Such morphological switching has been observed in serotypes A, B and D and. In *C. gattii* (serotype B) strains, cells can switch reversibly in the murine host, with mucoid variants again significantly more virulent than smooth variants. However, only smooth variants (with thinner capsules) are isolated from brain homogenate.[101] It has been demonstrated that phenotypic switching may be an important factor in relapse disease [123].

Another type of morphological change is the development of pseudohyphae (Figure 1-9), first detected when a subpopulation of *C. neoformans* with atypical forms survived incubation with *Acanthamoeba polyphaga*. However, these pseudohyphal strains were hypovirulent in mouse model [124]. A recent study dissected pseudohyphal strains (serotype A and D) derived from *in vitro* and exposure to amoebae. All of them possessed a premature stop codon in the *TAO3* gene, a component of the RAM (Regulation of Ace2p activity and cellular Morphogenesis) pathway. This pathway regulates cell type: it promotes polarity in the ascomycete *S. cerevisiae* but inhibits it in *C. neoformans* [36].

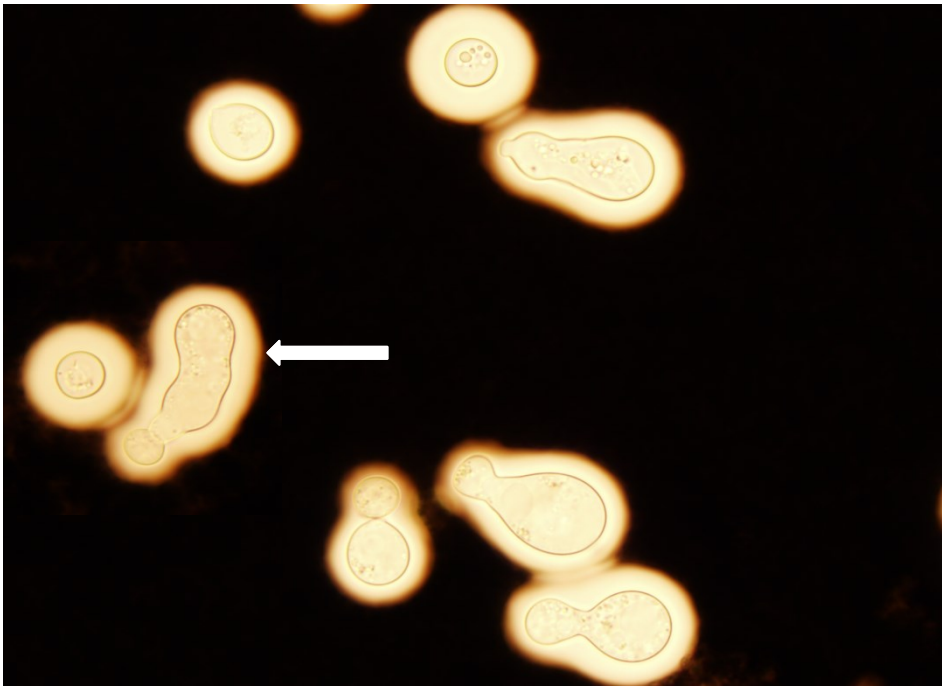


Figure 1-9 Pseudohyphal forms of environmental *C. neoformans* var. *grubii* were obtained after inoculation in *Galleria mellonella* for 48 hours at 37^{o}C. White arrow indicates a pseudophypha.

Another form of morphological transformation is enlargement (both body and capsule). This is believed to protect against host phagocytosis [125] [126]. These enlarged, so-called titan cells, which can be recovered from infected mouse lungs, reach diameters of up to 40-50 μm compared to the 5-10 μm of normal cells[127]. The proportion of titan cells can vary from 10 % to 80 % of the total cell count in the lung [127]. Titan cells can be induced *in vitro* in minimal media although the increase in size is less profound than seen *in vivo*. Titan cells are polyploid (e.g. tetraploid or octoploid) and reach their size by repeatedly replicating DNA without fission. RIM101, pheromone and cAMP signalling pathways are involved in regulating this giantism [127] [128]. Surprisingly, titan daughter cells have normal size but retain the same resistance to oxidative and nitrosative stress as their parental giant cells[128]. Thus, the enhanced survival of titan cell progeny may confer an advantage in

dissemination and persistence of pulmonary infection. Taken together, morphogenesis is important during invasion and evasion from the host immune response.

1.4.7 Degradative enzymes

1.4.7.1 Protease

C. neoformans exhibits extracellular proteolytic activities to degrade host proteins, such as collagen, elastin, fibrinogen, immunoglobulins, and complement factors, utilizing them as sources of carbon and nitrogen [129]. *C. neoformans* also employs protease and phospholipase to degrade phagosomal membrane during intracellular reproduction. [102]

1.4.7.2 Phospholipase

This is a heterogenous group of enzymes whose role is to hydrolyse ester linkage of phospholipase, thereby destabilising cell membrane and leading to cell lysis. Phospholipase exhibits not only hydrolase activity of phospholipase B (PLB) but also lysophospholipase, and transacylase as well. PLB mediates destabilization of the phagosomal membrane, facilitating fungal access to nutrients in macrophage cytoplasma [32]. PLB also takes part in antiphagocytic activity by cleaving protozoan- or mammalian- derived phospholipids into polar heads, which are sensed *C. neoformans* cells, inducing capsular enlargement and formation of giant cells to evade phagocytosis [130]. PLB can enhance adhesion of *C. neoformans* to human lung epithelial cells and contribute to dissemination of cryptococcosis from the lung to the blood via mononuclear phagocytes in murine model [15]. The *PLB1* mutant (PLB deficient) showed remarkedly reduced virulence compared with the wild-type strain

in the murine and rabbit infection models [131]. The *PLB1*-reconstituted strain can restore its PLB activities.

1.4.7.3 Urease

Urease catalyzes the hydrolysis of urea to ammonia and carbamate. Urease has been shown to enhance fungal dissemination to the brain [15]. Mice infected with urease-active strains have higher fungal burdens in the brain and higher mortality than those infected with urease-defective strains [132]. Urease gene (*URE1*)–deficient *C. neoformans* isolates displayed increased survival of infected mice compared to the wild type strain H99 and the avirulent phenotype can be restored by complementing *URE1* [133]. A urease-negative strain was isolated from CSF of an immunocompetent patient [134].Interestingly, this strain can restore urease activity and virulence in mice comparable to that of H99 strain by complementation of *URE2* gene. *URE1* failed to complement the mutant phenotype in this case. This suggests urease is dispensable for virulence of *C. neoformans* if other virulence factors can compensate for it. Experimentally, urease activity is used to detect presumptive cryptococci [135].

1.5 Signal transduction pathways regulating pathogenicity

To survive saprophytic and parasitic life in nature and the animal host, *Cryptococcus* deploys virulence attributes in resistance to environmental stresses: oxidative burst, phagocytosis, nutrient depletion, radiation exposure, desiccation. To regulate and orchestrate virulence factors, cryptococci have to employ complex signaling networks. There are 5 or 6 signaling pathways involved in modulating morphological switching, virulence and stress response. These key regulatory pathways involved

with virulence include cAMP/Protein kinase A, MAPK, Ca^{2+}/Calcineurin and Ras pathways.

1.5.1 Cyclic AMP/Protein kinase A pathway (cAMP-PKA)

This pathway regulates morphology and virulence including capsule formation, melanin production and mating. The pathway components are illustrated in Figure 1-10. Upstream components are G protein coupled receptor (GPR4) and Gα subunit (GPA1)[136]. Adenyl cyclase-associated protein ACA1 functions in parallel with GPA1 to control CAC1[137]. These proteins together with HCO_3^-, produced from CO_2 by carbonic anhydrase (CAN2), regulate activity of membrane-bound adenylyl cyclase (CAC1), thereby generating cAMP for downstream activation. Interestingly, activity of catalytic subunits of PKA is serotype dependent [138]. In serotype A, *PKA1*-deficient strains are avirulent with defects in capsule size, melanin formation and mating while PKA2 is not required for virulence. In serotype D, *PKA2* mutant exhibits defects in synthesis of melanin and capsule. The importance of cAMP-PKA in governing virulence of *C. neoformans* is demonstrated by finding the disruption of pathway components (GPA1, GPR4 and CAC1) leads to sterility, and defects in melanin and capsule formation, thereby reducing virulence in the mouse model of cryptococcosis [139]. In contrast, knockout of *PKR1*, gene encoding protein kinase A regulatory subunit, induced enlarged capsule and hypervirulence. A recent transcriptional comparison of mutants (*PKA1 & PKR1*) and wild type strain revealed differential expression of genes encoding for secretory pathway components, such as translocation (Sec61 and Hsp19/Kar2), vesicle formation and fusion (Bet1, syntaxin), Golgi transport (a-1,6-mannosyltransferase) and genes encoding chaperones, stress,

and phospholipid metabolism[140]. The role of PKA in transporting secretory components suggests PKA regulates expression of secretory pathway constituents to control expression of virulence traits at the cell surface. Moreover, *GPA1, CAC1* and *PKA1* mutants are more susceptible to Amphotericin B than wild type strains do, suggesting a role of the cAMP pathway in regulating resistance to polyene drugs [141]. Deletion of *GPA1*, *CAC1, ACA1, PKA1* also result in sensitivity to osmotic stress (caused by NaCl and KCl). Thus, taken together, cAMP-PKA pathway is a potential target for anticryptococcal therapy.

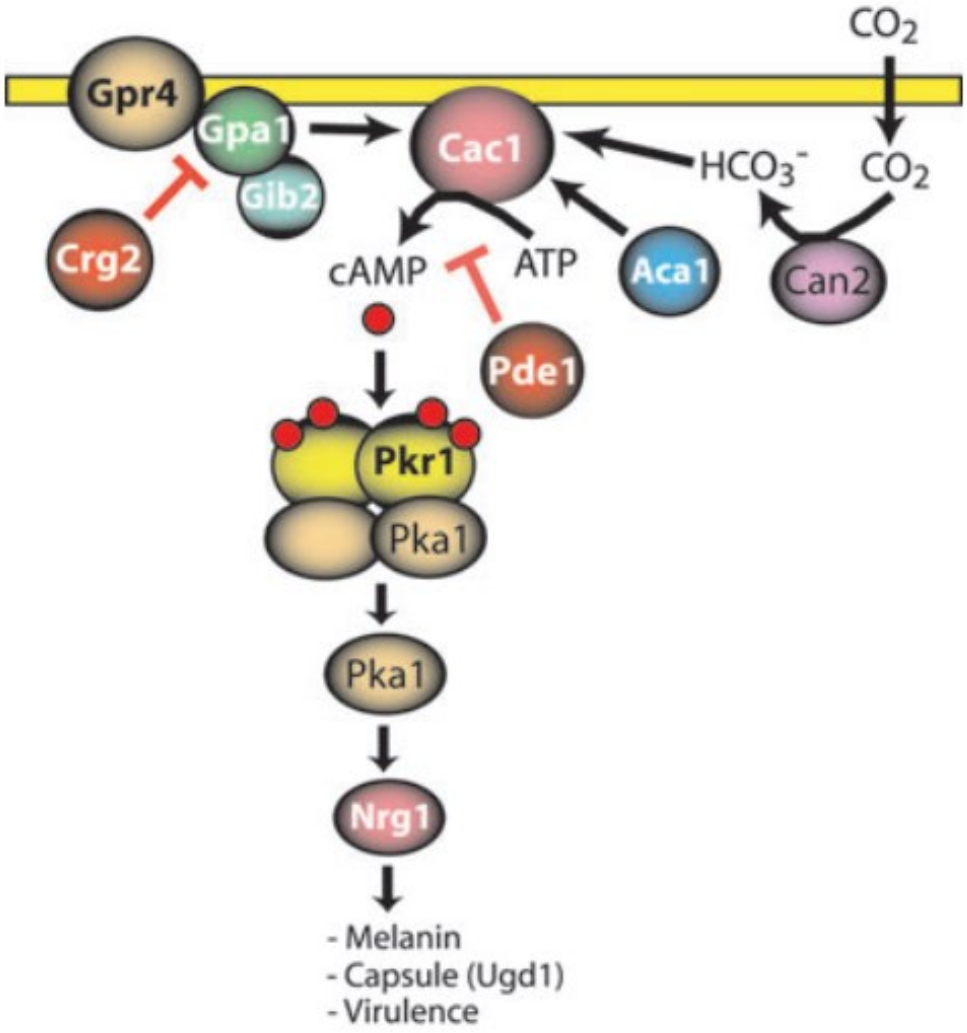


Figure 1-10 cAMP-PKA signal transduction pathway in the pathogenesis of *C. neoformans*. cAMP signaling responds to environmental cues (glucose & amino acid) through G-protein coupled receptor (GPR4) and G subunit or HCO3-. Once activated by GPA1, enzyme adenylyl cyclase (CAC1) releases cAMP, which subsequently activates Protein kinase A (PKA) by dissociating regulatory subunit Pkr1 from the two catalytic subunits of PKA (PKA1 &PKA2). PKA1 activates downstream proteins such as Nrg1. A GTPase-activating protein (Crg2) and phosphodiesterase (Pde1) negatively regulate Gpa1 and production of cAMP. Adapted from Kozubowski et al[138]

1.5.2 Mitogen-activated protein kinase (MAPK) pathways

Eukaryotic cells employ this pathway to sense environmental changes. Cascade components differ between higher mammalian and lower eukaryotic cells, but the core of MAPK pathways is conserved[142]. The core of pathways consists of three

kinases. Upon environmental stimuli, MAP kinase kinase kinase (MAPKKK) phosphorylates MAP kinase kinase (MAPKK), which in turn triggers phosphorylation of MAP kinase (MAPK). The signal is eventually relayed to downstream effectors, activating or suppressing the corresponding target genes (Figure 1-11).

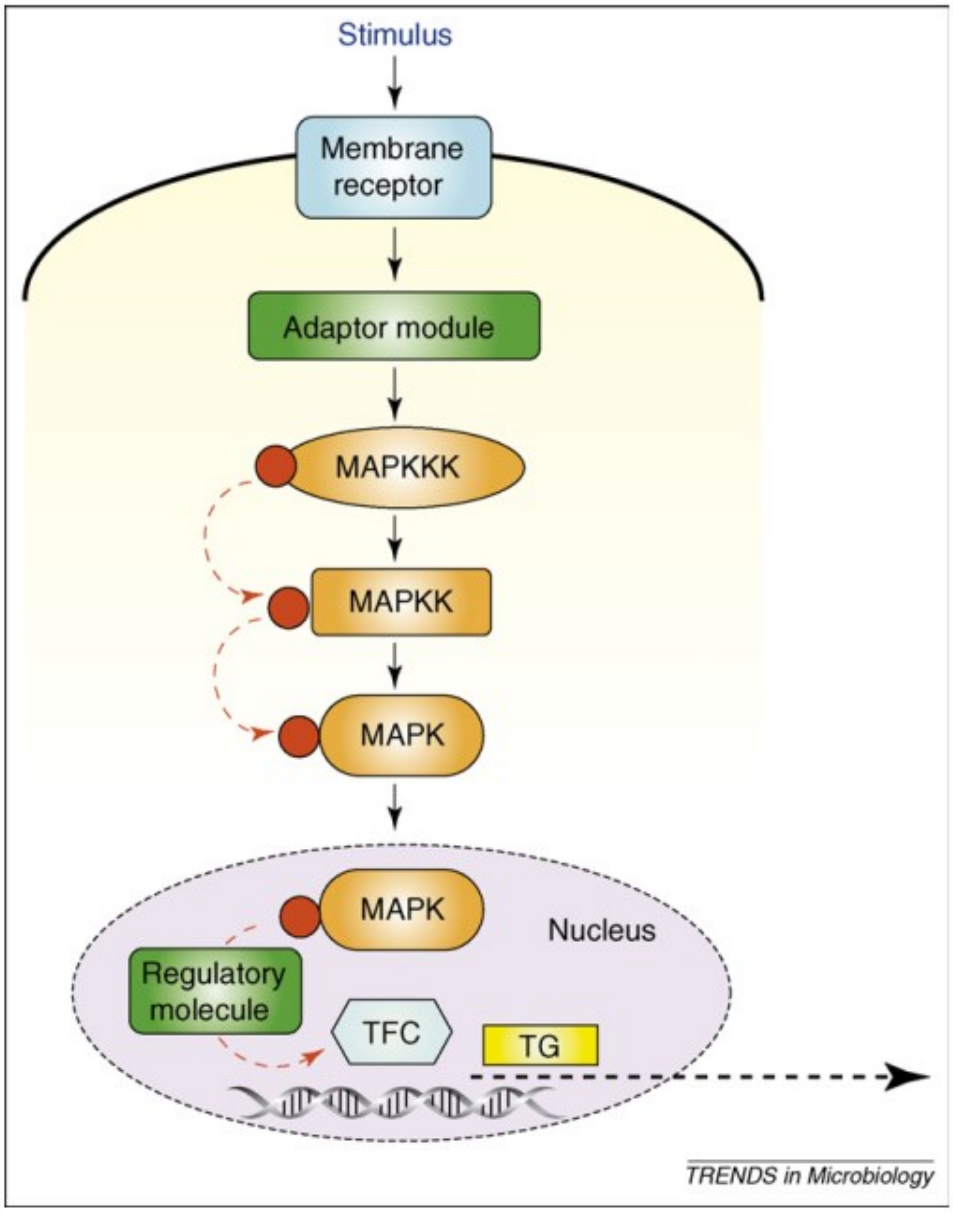


Figure 1-11 The general scheme of the MAPK cascade. Stimuli are perceived at the plasma membrane through membrane receptor. The signal is relayed by adaptor module to the core component of the cascade through phosphorylation. The MAPK pathway consists of 3 kinases: MAPKKK, MAPKK and MAPK. Upon phosphorylation, MAPK is transported to nucleus where it phosphorylates a transcription factor, which switches on or off the downstream target genes, followed by occurrence of adaptive response to extracellular cue. Adapted from Roman *et al*.[142]

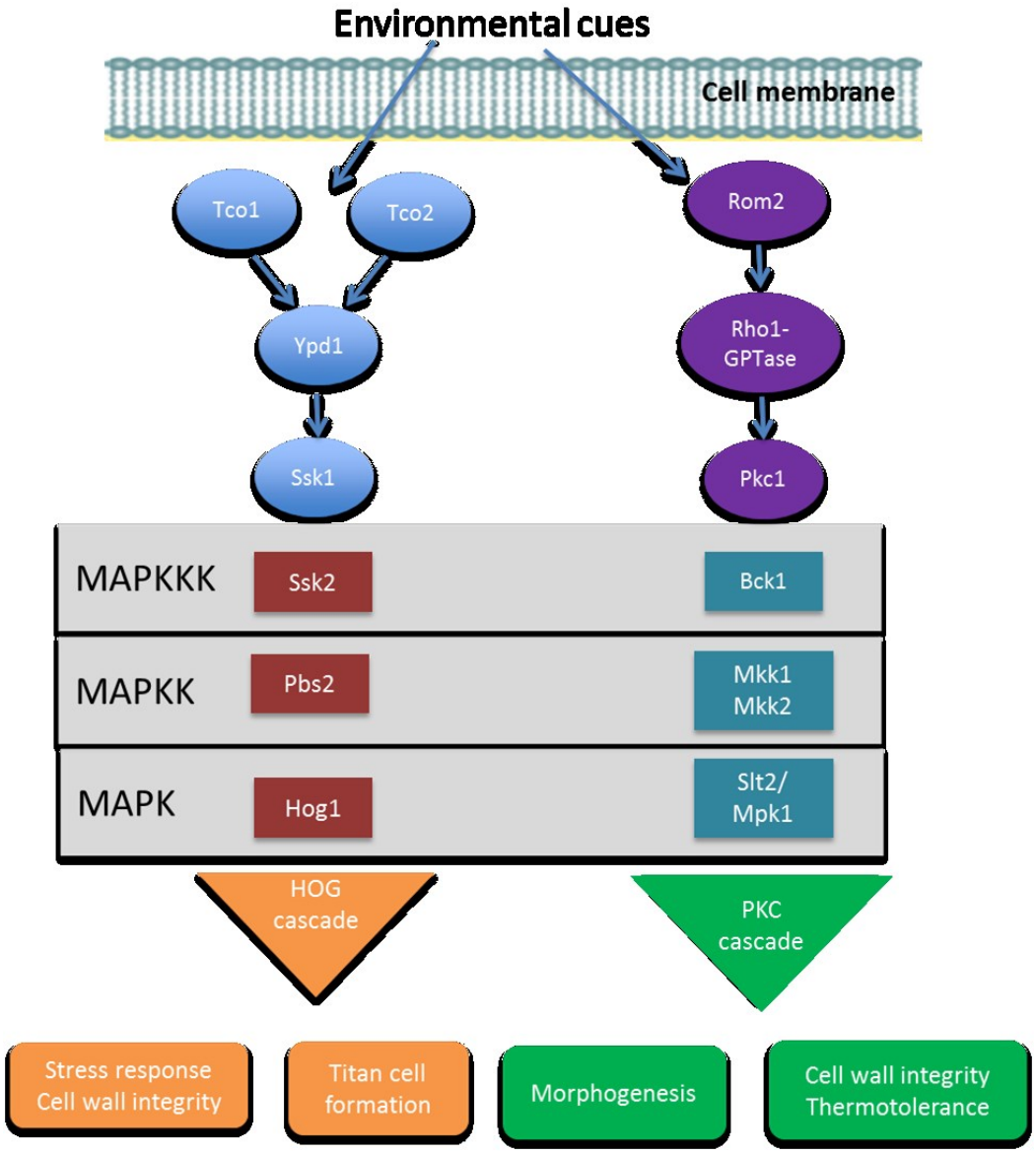


Figure 1-12 Components of MAPK signaling pathways including HOG pathway and Protein kinase C (PKC) pathway and their roles in regulating cryptococcal virulence determinants. Adapted from O'Meara and Alspaugh *et al*[143]

In *S. cerevisiae*, at least five MAPK pathways regulate response to shocks (oxidative, nitrosative, osmotic) and pheromone during mating. MAPK is also crucial for cryptococcal survival during mammalian infection. Sensing and adaptation to environmental stresses in *C. neoformans* serotype A are exclusively mediated by the high-osmolarity glycerol (HOG) cascade. The signal transduction cascade is initiated when two sensor kinases (Tco1 &Tco2) are stimulated by environmental cues [144].

Once phosphorylated, Ypd1 histidine kinase relays signal to Skk1 response regulator. Under normal growth Hog1 MAPK is constitutively phosphorylated (inactivated form) by Pbs2, negatively inducing capsule and melanization (Figure 1-12). In response to stresses, Tco1 and Tco2 activate Ssk1, which in turn triggers phosphatase to dephosphorylate Hog1 MAPK for activation of downstream effectors to counteract stresses. Ssk1 is critical response regulator. Ssk1 mutation lead to enhanced production of melanin and capsule and defect in Hog1 MAPK constitutive phosphorylation in normal conditions [144]. There are numerous kinase and transcription factors downstream of Hog1 that are still understudied. Hog1 interferes with cAMP-PKA pathway, repressing formation of melanin and capsule in serotype A. Hog1 mutants produce larger capsules than wild type (H99) and restore melanin production in GPA1, CAC1 and PKA1 mutants [145]. This provides evidence that there is cross-talk between cAMP-PKA and HOG pathway. In serotype A, Hog1 is constitutively phosphorylated under normal growth but rapidly dephosphorylated in response to osmotic stress. However, this occurs conversely in serotype D. Hog1 protein sequence is identical between serotype A and D but phosphorylation pattern differs, suggesting upstream components of Hog1 are divergent. *Hog1* and *Pbs2* mutants exhibit hypersensitivity to high temperature (40^{o}C), UV radiation and osmotic stress. Disruption of Hog1 or Pbs2 lead to attenuated virulence in mice although both mutants significantly enhance production of melanin and capsule, which confirm the mutifactorial nature of *C. neoformans* virulence and underscore that the ability to cope with stress, capsule formation and melanization are important in pathogenesis.

PKC (Protein kinase C) cascade is another stress-activated pathway essential for cell wall integrity and thermotolerance (Figure 1-11). Pkc1 is also phosphorylated in response to reactive oxygen and nitrogen species but Pkc1 downstream components (Bck1 and Mkk2) have no role in protection against these stresses. Pkc1 mutants show growth defects upon exposure to 37^{o}C incubation, oxidative and nitrosative stresses, cell wall inhibitors and additionally reduce melanization. Pkc1 mutants cause more severe defects in cell wall biogenesis than *Bck1* and *Mkk2* mutants, implying an additional role in cell wall biogenesis beyond activation of the cascade [138]. Taken together, Pkc1 is a central coordinator with multiple outputs other than activating the MAPK cell integrity cascade [146]. Phosphorylation of Pkc1 initiates a transduction cascade involving downstream components: Bck1, Mkk1, Mkk2 and Mpk1. Bck1 and Mkk2 mutants are critical for cell wall integrity. They exhibit increased susceptibility to cell wall inhibitors (SDS and Congo red) [147]. Mpk1, the terminal component of this pathway, is important for growth at 37^{o}C in vitro, which, as discussed earlier, is a prerequisite for human infection. Thus, Mpk1 mutants display attenuated virulence in mouse [148]. Mpk1 is phosphorylated following exposure to cell wall defect-induced antifungals (inhibitors of chitin synthase and b-1,3-glucan synthase), oxidative and nitrosative stresses. A recent study demonstrated an interconnection between the HOG and PCK pathway via casein kinase I protein (Cck1). Cck1 regulates cell wall integrity and adaptation to osmotic shock through phosphorylation of Mpk1 and Hog1 proteins [149].

1.5.3 Ca2+/Calcineurin pathway

This Ca^{2+} mediated signaling pathway takes part in diverse functions, including cell wall integrity, sexual differentiation, thermotolerance, and adaptation to cell wall stress, high CO_2 and alkaline pH [150]. In response to environmental signals, a rise in levels of cytosolic Ca is sensed by Calmodulin (Cam1), which can then bind four Ca^{2+} molecules to form the Ca^{2+}/calmodulin complex (Figure 1-13). This complex subsequently activates Ca/calmodulin-dependent protein kinase (CaMK) or calcineurin, a serine/threonine phosphatase. Calcineurin comprises two subunits: catalytic A subunit (Cna1) and regulatory B subunit (Cnb1). Upon activation, calcineurin triggers transcription of downstream genes which protect against environmental stresses. Activation of calcineurin can be inhibited by antifungals/immunosuppressive drug, such as Cyclosporine A (CsA), tacrolimus (FK506) and tamoxifen.

These substances form complexes with Cyclophilin A and FK506-binding protein (FKBP12), respectively. The complexes inhibit growth at 37^{o}C but not 24^{o}C [151]. Calmodulin can act independently of Ca^{2+}/calcineurin to regulate morphogenesis and growth at 37^{o}C [152]. Calcineurin is essential for survival at 37^{o}C. The *CNA1* mutant suppresses growth at elevated temperature in vitro and was avirulent in the rabbit model of cryptococcosis [151]. To survive intracellularly, *Cryptococcus* must adapt to increased levels of CO_2 and physiological pH. *CNA1* mutants display poorer growth at 5% CO_2 and pH 7.3 in RPMI broth compared with wild type strains. Disruption of Cnb1, the calcineurin regulator subunit, also impairs growth at 37^{o}C and renders strains avirulent in the murine model[153]. Moreover, calcineurin deficiency (*CNB1*

deleted) also induces phosphorylation of Mpk1 MAPK and upregulation of the *FKS1* gene, which encodes a subunit of b-1,3-glucan synthase. This indicates that calcineurin and Mpk1 coordinately regulate cell wall integrity in response to increased temperature stress [148]. Patients are believed to acquire cryptococcosis by inhaling infectious spores from these environments. Finally, calcineurin is involved in the mating process – drugs which inhibit it (Cyclosporin A and FK506) inhibit mating and haploid fruiting [154].

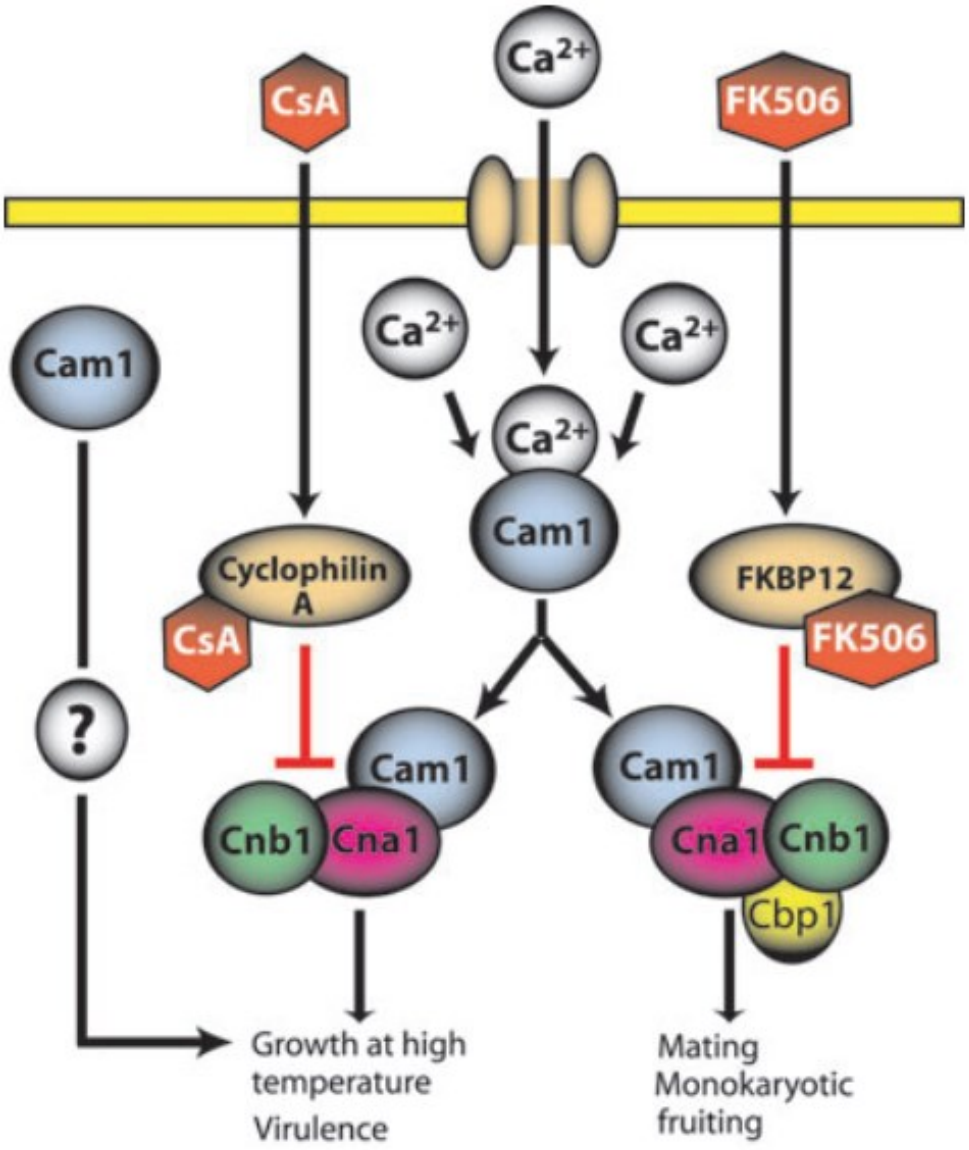


Figure 1-13 Ca2+/calcineurin signaling cascade and their roles in regulating virulence of *C. neoformans*. Cyclosporine A (CsA) and FK506, bound to cyclophilin A and FKBP12, respectively, inhibit calcineurin. Cbp1: calcipressin, a calcineurin effector. Adapted from Kozubowski *et al* [138]

1.5.4 Ras pathway

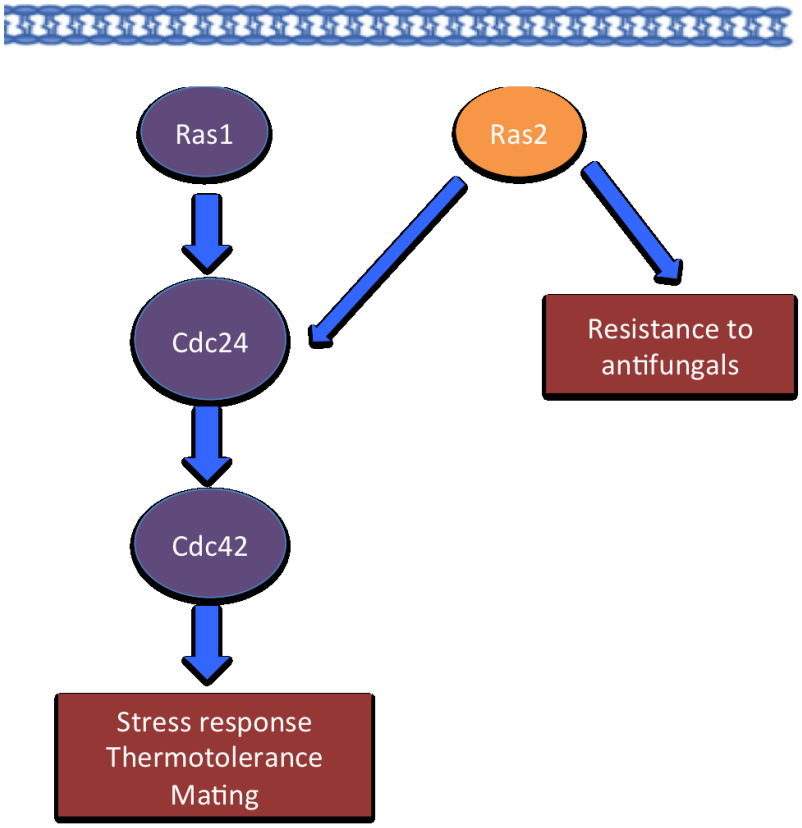


Figure 1-14 Ras signaling cascade and its role in regulating thermotolerance, stress response and pheromone response during mating. Adapted from Shinae Maeng *et al*[155]

The Ras signal transduction cascade is also a stress-triggered signaling pathway. It governs growth at elevated temperature and adaptations to environmental stress and pheromone response in mating (Figure 1-14). Ras proteins are guanine binding proteins. They are active in the GTP-bound from and inactive in GDP-bound form following GTP hydrolysis. *RAS1* mutants display growth defects at 37^{o}C and an avirulent phenotype in the rabbit model of cryptococcosis. Mating filaments were not observed when *RAS1* mutants were incubated with opposite mating type strains, unlike the Ras1 wild type strains. However, defective mating of *RAS1* was partially suppressed upon introduction of cAMP [156], implying potential cross-talk between Ras

and the cAMP signaling pathway. *RAS1* has been reported to induce expression of pheromone response pathway genes: α-pheromone gene *MFα1* and α- pheromone receptor gene *CPR1α* [157]. *RAS2* is expressed at much lower level than *RAS1* and is not required for pathogenicity of *C. neoformans*. *RAS2* mutants can grow well at elevated temperature, exhibit normal mating and are virulent comparable to wild type strains in the mouse model. However, *RAS1 RAS2* double mutants display more severe growth defects at high temperatures than strains with either single *RAS* gene deletion, suggesting interplay of these two genes in regulating response to high temperature. Overexpression of *RAS2* in *RAS1* mutant background partially suppresses growth defect at elevated temperature and completely suppressed mating defects of *RAS1* mutant strains, implying an overlapping regulatory function [158]. A recent study found Cdc24, guanine nucleotide exchange factor, functions downstream of Ras1 and upstream of Rho-like GTPase Cdc42 to mediate high temperature growth and pathogenicity [159]. Besides its role in adaptation to elevated temperature, *RAS1* mutants exhibited sensitivity to hyperosmotic stress and cell wall-destabilizing agents, indicating importance of Ras1 pathway for cell wall integrity. Both *RAS1* and *RAS2* mutants are more susceptible to Amphotericin B than wild type strains. Additionally, disruption of these genes did not change significant expression of ergosterol synthesis genes. This evidence implies Ras proteins can regulate cryptococcal resistance to AMB without interference with ergosterol production [141]. Given Ras and cAMP/PKA signaling pathway involvement in susceptibility to stress response and AMB, their components are potential targets for discovery of potent antifungal therapy.

1.6 *Galleria mellonella* infection model

In vitro and *in vivo* infection models have been used to investigate pathogenicity of *Cryptococcus* pathogens. Many cryptococcal virulence traits are only induced in hosts because *in vitro* models cannot completely mimic the stresses that cryptococci encounter *in vivo*, such as phagocytosis, nutrient deprivation, elevated temperature, etc. The mammalian model most frequently used is intranasal inoculation of mice (*Mus musculus*), which is expected to simulate the naturally occurring infection process [160]. However, mammalian models using mice, rats and rabbits are costly, have small brood size, are time-consuming and have significant ethical considerations.

The finding that pathogenesis attributes involved in cryptococcal virulence remains similar in both vertebrate and invertebrate hosts make the latter appealing for study of fungal pathogenesis in mammalian hosts. Genes associated with mamalian virulence (*CAP59, GPA1, PKA1, RAS1, PKA1*) have also been demonstrated to be significant virulence factors in invertebrate models including *Caenorhabditis elegans, Drosophila melanogaster and Galleria mellonella* [98] [161].

The larval stage of the greater wax moth *Galleria mellonella* (Figure 1-15) has been extensively employed as a model of the innate immune response to study the pathogenesis of a number of microbial pathogens, from bacteria to fungi: *Bacillus cereus, Acinetobacter baumannii, Listeria monocytogenes, Staphylococcus aureus, Candida albicans, Aspergillus fumigatus, Aspergillus flavus, C. neoformans* [162]. The innate immune response is a critical component of the mammalian immune system response to infection, and it has been shown that comparable data are obtained

when both the murine and wax moth larval model are used to study pathogenesis [163–165]. Moreover, larvae are cheaper to purchase, easier to maintain, thermotolerant and are not subjected to the same ethical issues as mammals. Results are available within two weeks. Therefore, they are used to screen a large number of isolates for virulence or to assess efficacy of antifungal/antimicrobials. The selection of the *G. mellonella* model depends on the research question and innate immune response of interest. However, if the question pertains to organ dissemination or organ-specific fungal burden (brain, lung, spleen), then the *G. mellonella* is not appropriate simply because the larvae do not have comparable structures to such organs.

Figure 1-15 Larvae of the greater wax moth *Galleria mellonella*

The immune system of *G. mellonella* possesses a number of functional and structural homology to the innate immune system of vertebrates, including humoral and cellular defenses [162]. The humoral immune response mediates larval melanization, hemolymph clotting and antimicrobial peptides while the cellular response mediates phagocytosis and nodulation (encapsulation of pathogens within layers of

hemocytes). Larval hemolymph functions similarly to blood in higher animals. It transports nutrients, waste and immune cells called hemocytes. At least six types of hemocytes have been identified which function similarly to human neutrophils through phagocytosis and generation of superoxide to protect against pathogens [165].

The *G. mellonella* model has been used to test efficacy of antifungals (Amphotericin B, Fluconazole, Flucytosine when used alone or in combination [98]. The combination of Amphotericin B and Flucytosine is more effective than Amphotericin B alone in this model. The *G. mellonella* model has been shown to delineate the major molecular types of *C. gattii* strains (VGI, VGIIa, VGIIb, VGIII and VGIV) [166]. In another study, titan cells were observed in larvae infected with *C. neoformans,* and *G. mellonella* extract can induce capsule enlargement when cultured with *C. neoformans in vitro* [167]. One weakness of the *G. mellonella* model so far is that standardised conditions for maintenance and reproduction of larvae have not been established yet, which could make inter-laboratory comparisons difficult. A recent study demonstrated that larvae deprived of food 7 days prior to infection were more susceptible to *C. albicans* than unstarved larvae. Proteome analysis of hemolymph from starved larvae shown reduced expression of antimicrobial peptides and immune proteins, which is likely to enhance susceptibility of larvae to infection .[168]

1.7 Rationale and importance of studies in this thesis

Growing evidence underlines *C. neoformans* var. *grubii* as a prominent pathogen causing infection in both immunocompetent and immunosuppressed patients [1 86 87]. In Vietnam, Day *et al*. identified an AFLP-defined cluster of VNI they termed VNIγ (later identified to be MLST sequence type 5, and genomic lineage VNIa-5 (Ashton et

al, https://www.biorxiv.org/content/biorxiv/early/2018/06/28/356816.full.pdf) responsible for the majority of infection (>80%) in apparently immunocompetent patients[86] [88]. Therefore, I hypothesize that the ST5 lineage has increased pathogenicity compared with other *Cryptoccocus neoformans* strains. The HIV infection prevalence rate in Vietnam is relatively low at approximately 1% of adults. The fact that ST5 isolates only account for ~35% of disease in the HIV infected population suggests it is not the dominant strain circulating in the environment, if they truly have increased pathogenicity. The lineage ST5 is of great interest because it also predominates in China and neighbouring countries in north eastern Asia (Japan, Korea) where disease in HIV-uninfected patients has been frequently reported[169].

The ability to cause human disease is likely a by-product of adaptation to a particular environmental niche, which accidentally confers pathogenicity. I hypothesize that the proposed increased pathogenicity of the ST5 isolates is a result of adaptation to a particular ecological niche. Identifying the ecological niches of environmentally required cryptococcal pathogens could enable the management of potential cryptococcosis outbreaks. Associations between outbreaks and ecological niches of *Cryptococcus* have been established in endemic areas around the world, with the identification of native trees and avian droppings as sources of infectious propagules [26,52,53,57]. However, the veracity of these findings is undermined by the fact that environmental sampling has never been randomised, and thus the results of environmental sampling cannot be assumed to be representative of the true environmental population. A comprehensive and randomised sampling approach to

interrogating the ecology of *Cryptococcus* is warranted, to better understand both the prevalences of different lineages, and to accurately describe the ecology. Understanding cryptococcal adaptation to ecological niches may provide insight into how pathogenic forms emerge.

The globally predominant strain is the VNI, serotype A lineage, which is isolated from both clinical and environmental settings. Person to person spread is not considered to be a method of disease acquisition; therefore, competent pathogenic isolates are definitely present in the environment. However, there are discrepancies in virulence between clinical and environmental isolates, this having been described in the USA (North Carolina, California), Brazil, and Puerto Rico. It is not clear whether this discrepancy relates to microevolution within major lineages, or is some epigenetic effect. The wide dispersal of strains from HIV uninfected patients throughout the VNIa-5 lineage (Figure 1-6) suggests that all strains within this lineage have the potential to be highly pathogenic (cause disease in immunocompetent patients). The cause of variability in pathogenicity of strains identical in genotype but from different sources is intriguing. Understanding the mechanism that defines this variability could lead to the discovery of novel drug targets.

1.8 Aims of the thesis

a. To determine whether there is temporal-spatial clustering of *C. neoformans* by genetic lineage and HIV status in Vietnam

b. To identify ecological niches associated with these lineages;

c. To define the diversity of the Vietnamese *C. neoformans* population in the environment and prevalence of particular lineages;

d. To characterise the virulence phenotypes of environmental *Cryptococcus* isolates;

e. To compare the virulence between and within different genotypes and sources using the *G. mellonella* model;

f. To define the transcriptional differences associated with variability in virulence

g. To identify gene candidates associated with variability in virulence between isolates.

Chapter 2

Spatial-temporal trends in cryptococcal meningitis at the Hospital for Tropical Diseases, Ho Chi Minh city

2.1 Introduction

The estimates of the global burden of HIV-associated cryptococcal meningitis were revised in 2014 at 230,000 cases per year[1]. This human epidemic has driven a number of efforts to identify the spatial distribution and ecological niche of the responsible pathogens *C. neoformans* and *C. gattii*. The ecological niche of a species is defined as an area with favourable physical and biological conditions within which the species can survive and reproduce to maintain its population, without the need for immigration[170]. Understanding the ecological niche of a pathogen should give insight into the evolution of its pathogenesis and might allow control or interventions. For *C. neoformans*, the association with pigeon excreta was identified in North Carolina, USA by Litvintseva *et al* in 2004[171]. The importance of trees and the link with the eucalyptus tree was identified by Ellis *et al* in Australia in 1990[50]. Since these seminal studies, there have been a number of efforts to identify the ecological niches of these species elsewhere. An association with avian guano has been identified in birds, including in southeast Asia, where (dried) chicken guano was an found to be a source in suburban areas of Phayao province, northern Thailand[57]. In South America, environmental sampling in Columbia identified *C. neoformans* serotype A to be the most prevalent type, followed by *C. gattii* serotypes B and C[172]. Here, the *Cryptococcus gattii* VGIII molecular type (serotype C) was associated with almond trees (*Terminalia catappa*)[172]. In Zambia, Africa, Fisher *et al* found adaptation of environmental *C. neoformans* isolates lineage VNB (87%) were almost found in rural areas dominated by mopane trees while the majority of clinical isolates (76%) belonged to lineage VNI [173].

However, until the *C. gattii* outbreak in British Columbia, there has been little evidence of clustering of human disease. Driven by interest in the Canadian outbreak of *C. gattii* disease, Kidd *et al* investigated ecological determinants of *C. gattii*

colonization in British Columbia and the northwestern United States. They identified colonization hot spots which consisted of acidic soil associated with low moisture and low organic carbon content[53]. In Vietnam *C. neoformans* VNI is responsible for almost 100% of cryptococcosis in HIV infected patients at the Hospital for Tropical Diseases (HTD), Ho Chi Minh City. Disease also occurs in HIV uninfected patients, who usually do not appear to have associated underlying immune deficit. *Cryptococcus gattii* is responsible for about 20% of disease in these patients, and the rest are due to *Cryptococcus neoformans* molecular genotype VNI. Using AFLP, we previously defined 2 clusters of VNI circulating in Vietnam which we termed VNIϒ and VNIδ. VNIϒ strains caused 84 % of infection in HIV-uninfected patients, but only 38% of disease in HIV-infected patients (odds ratio, 8.30; 95% CI, 3.04 to 26.6; P < 0.0001)[86]. We subsequently used the consensus MLST scheme to show that the VNIϒ precisely corresponded to MLST ST5, and we hypothesize that the VNIϒ/ST5 cluster has increased pathogenicity compared to other VNI clades. Since *C. neoformans* is primarily an environmental saprophyte, this increased pathogenicity presumably represents an adaptation to a particular ecological niche which coincidentally confers increased pathogenicity.

Statistical methods have been widely used to determine spatial variation of infection outbreak. One type of variation analysis is clustering analysis which examines if geographical cluster of disease occurs by chance alone. Detection of temporal-spatial cluster was important for infectious disease surveillance, such as dengue, malaria, hand foot mouth disease, West Nile virus, Hanta virus, typhoid fever [174–179].

OUCRU and HTD have enrolled significant numbers of patients into cryptococcal meningitis trials. This represents an opportunity to identify whether there is clustering of cases, which should provide insights into the ecology of *C. neoformans* in the south of Vietnam. Understanding adaptation of pathogens to ecological niche will pave the way to understand how they emerge as pathogens. In this chapter I shall focus on a retrospective analysis of clinical isolates in Vietnam collected over 15 years through our trials.

2.2 Aims

My aims in this chapter are to:

1. Identify the geographical distribution and clustering of cryptococcal meningitis cases in southern Vietnam due to *C. neoformans seen* at the Hospital for Tropical Diseases and Cho Ray hospital over 15 years
2. Identify whether spatial-temporal clustering of clinical *C. neoformans* occurs by HIV status
3. Identify whether spatial-temporal clustering of clinical *C. neoformans* occurs by lineage
4. Identify whether there is clustering of *C. neoformans* sequence type 5 from HIV-infected patients for targeted environmental sampling in chapter 3

My analyses are limited to patients for whom I had viable *Cryptococcus neoformans* isolates.

2.3 Methods and Materials

2.3.1 Strains

I used data from cases of cryptococcal meningitis who had been enrolled in clinical trials at the Hospital for Tropical Diseases (HTD) and Cho Ray hospital (CRH), Ho Chi Minh City. The studies included a prospective descriptive study of cryptococcal meninigitis in HIV uninfected patients, a randomised controlled trial of antifungal therapy for cryptococcal meningitis in HIV infected patients, and a randomised controlled trial of adjunctive treatment for HIV associated cryptococcal meningitis[180] [67] [66]. A total of 409 *C. neoformans* isolates were available. All isolates were geocoded to level of the the patient's home address. All isolates had undergone whole genome sequencing within OUCRU and had been assigned to one of 7 major lineages: VNIa-4, VNIa-5, VNIa-32, VNIa-93, VNIa-outlier, VNIa-X and VNIb (manuscript under review, Philip Ashton *et al*). Of note, because no intervention trials were running for a period spanning 2011 into 2012, there is a break in recruitment of HIV-assocaited cryptococcosis patients during this period. The number of strains assigned to each lineage is presented in Table 2-1.

Lineage	Number of isolates	
	HIV-infected patient	HIV-uninfected patient
VNIa-4	167	12
VNIa-5	122	42
VNIa-32	15	4
VNIa-93	43	1
VNIa-outlier	1	
VNIa-X	1	
VNIb	1	

Table 2-1 Frequency of different genomic lineages of *C. neoformans* in the study

2.3.2 Spatial-temporal cluster analysis

The discrete Poisson model was used for the retrospective space-time scan statistic to detect spatial-temporal clusters with high rates of cryptococcosis per 100000 person-years[181]. Human population data was resolved to the district level for 15 years (2000 to 2014). The space-time scan statistic is a cylindrical window with a circular geographic base and height corresponding to the time period. The cylindrical window is then moved in space and time to scan for each possible geographical location and size at each possible time period. The relative risk is calculated inside and outside each window as well as log likelihood ratio. The window with the maximum likelihood constitutes the most likely cluster. Using Monte Carlo hypothesis testing, a P-value was obtained by comparing the rank of the maximum likelihood from the study data set with the maximum likelihoods from 999 simulations of the random data set.

The P-values are adjusted for the multiple testing due to multiple cylinders corresponding to different spatial and temporal locations and sizes of potential clusters. Secondary clusters were ordered according to their likelihood ratio test statistic and were reported if they did not geographically overlap with the most likely cluster. The cluster is significant if its simulated p-value was <0.05. The maximum spatial cluster size were set as <50% of the total population.

2.3.3 Covariate adjustment

In addition, I analysed clustering with adjustment for age and gender as covariates. These covariates were taken into account as categorical variables for the analysis[181]. I chose to adjust for these covariates because they are related to the disease. In Vietnam, cryptococcal meningitis is more frequently diagnosed in males compared with females. Age group is not randomly distributed. Young adults (26-36 years-old) are more prone to infection compared with other age groups (Table 2-2)

Gender	Number (%)
Male	**310 (76)**
Female	99 (24)
Age	
15-25	114 (28)
26-36	**206 (50)**
37-47	56 (14)
48-58	**18 (4.4)**
59-82	15 (3.6)

Table 2-2 Characteristics of 409 patients included in the analysis

2.3.3 Statistical analysis

Statistical analyses were performed using R version 3.1.2 (R Foundation for Statistical Computing, Vienna, Austria). Fisher's exact test was used to determine associations between ST5 and HIV-uninfected patients, and associations between ST5 and urban areas. The discrete Poisson model for spatial-temporal cluster analysis was performed using SatScan v9.4 (Maryland, USA). A P value of ≤0.05 was considered to be statistically significant.

2.3.4 Experimental workflow

Experiment workflow for this chapter is presented below

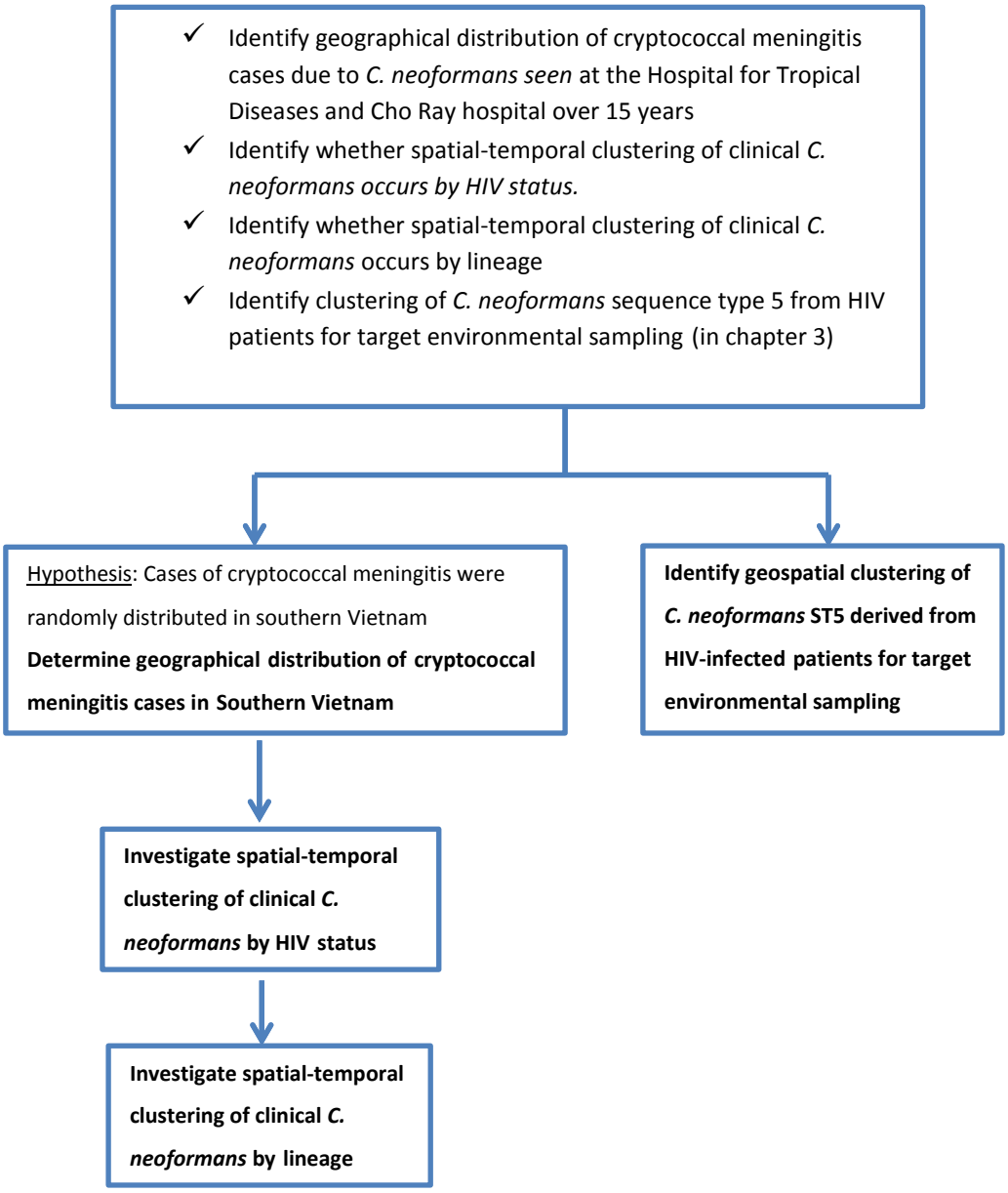


Figure 2-1 Experimental workflow of chapter 2

2.4 Results

2.4.1 Geographical distribution of *C. neoformans*

Over the past 15 years the number of cases of cryptococcal meningitis presenting to HTD and CRH each year has been distributed unevenly. The number of cases increased slowly from 2000, peaking in 2005 (Figure 2-2). This trend then decreases slowly when we did not recruit HIV patients for trials in 2011-2012. The trend reaches another peak in 2013 when the CryptoDex trial began recruitment.

Cases of cryptococcal meningitis seen at these hospitals had an uneven distribution from across the south of Vietnam, being distributed from Ca Mau at the southern tip of the country, to Hue in central region, with the highest number being from Ho Chi Minh city (207 cases, 51%) (Figures 2-3 & 2-4). Of 409 cases, 59 were isolated from HIV-uninfected patients. Lineage VNIa-5 was highly associated with HIV-uninfected patients (P value < 0.001, Fisher's exact test). HIV-uninfected patients travelled from as far north as Hue in central Vietnam to access treatment in Ho Chi Minh City. There were no cases in HIV infected patients from further north than Binh Dinh province.

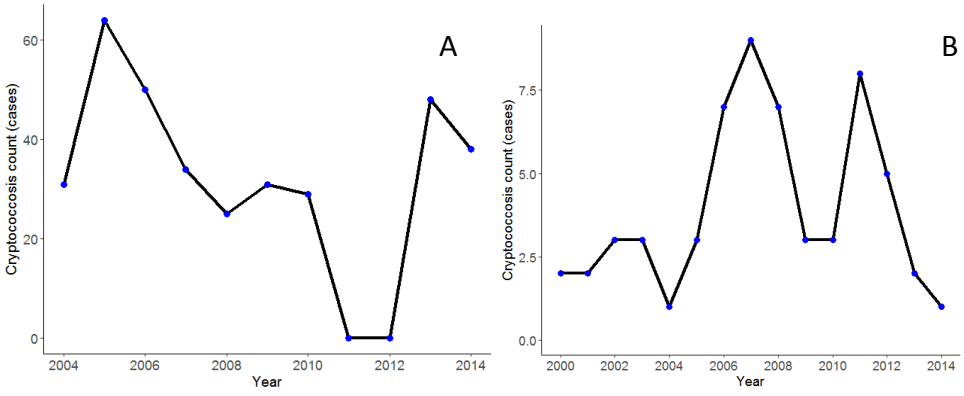


Figure 2-2 Longitudinal number of HIV-associated (A) and non-HIV cryptococcosis (B) cases in Vietnam from 2004 to 2014. Cryptococcosis count represents total number of cryptococcosis in Vietnam by year. There are no data for HIV–associated cryptococcal infection in 2011-2012 (we were not recruiting into a trial in those years).

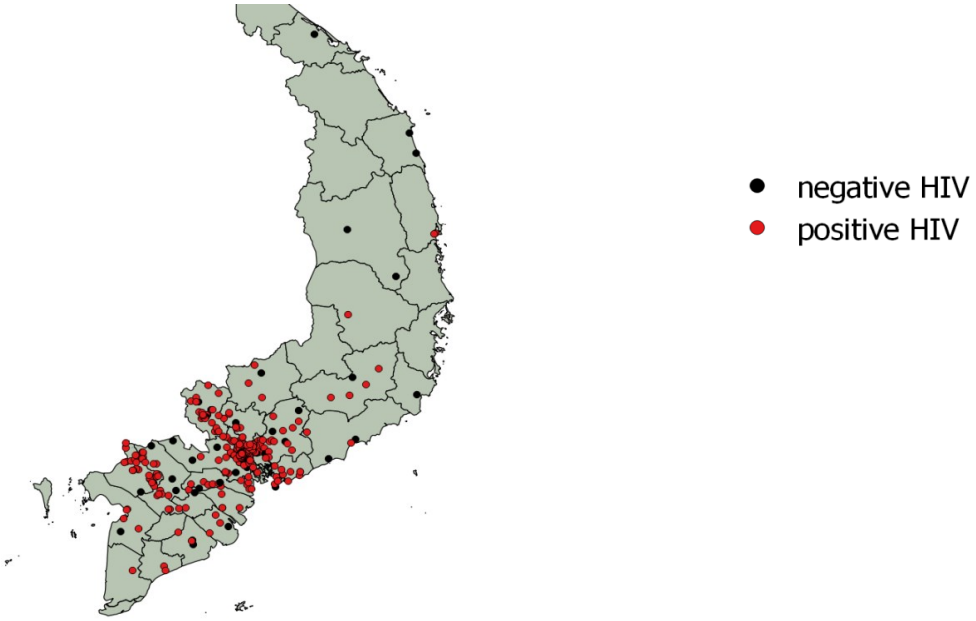


Figure 2-3 Geographical distribution of cryptococcal meningitis from HIV-infected and HIV-uninfected patients at the province level from 2000 to 2014. Each dot represents a single case.

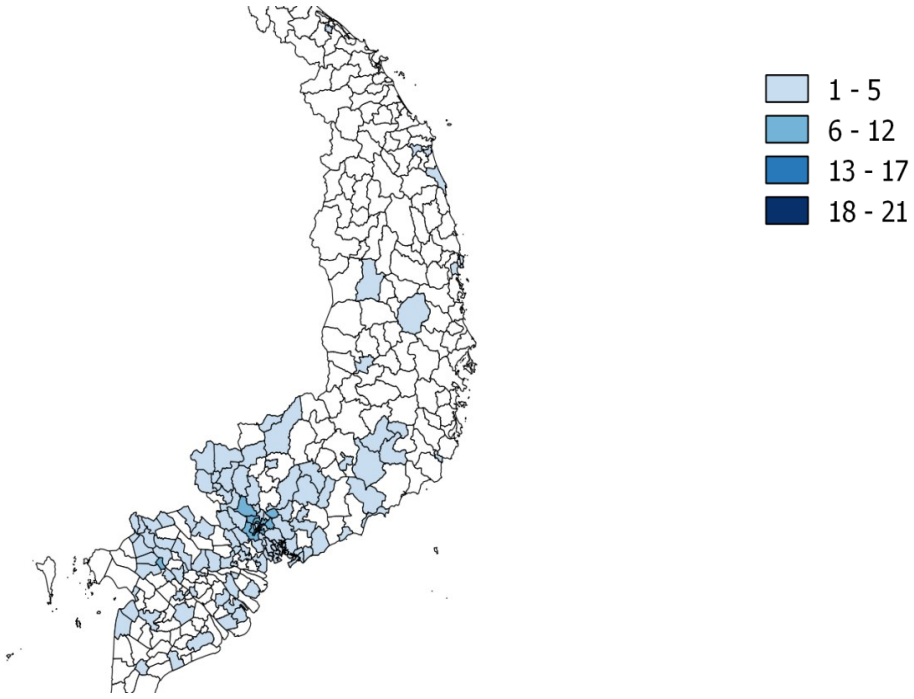


Figure 2-4 Number of cryptococcosis cases in Vietnam at district level from 2000 to 2014. The darker the colour is, the more cryptococcosis cases in each district.

The number of cryptococcosis cases in Ho Chi Minh city for 15 years was presented below (Figure 2-5). There are more cases from the north of the city, likely driven by the higher population density compared with the south (4898 people/km^2 vs 190 people/km^2). The majority of districts in southern HCMC are rural areas, covered by mangrove forest and sparsely inhabited. However, the highest burden of disease was in Binh Thanh district, which ranks only third for total population, and 12th in population density compared with other districts. In the Mekong delta, the largest number of cryptococcal meningitis cases came from An Giang province, which is the most populous.

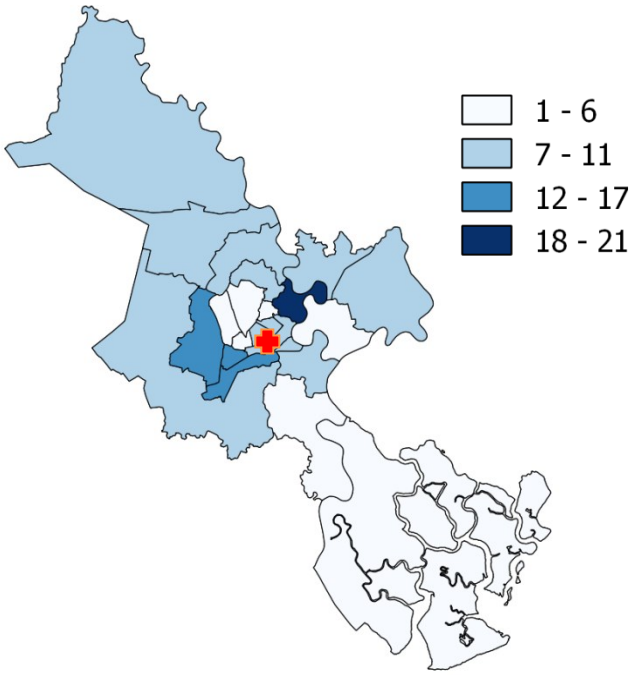
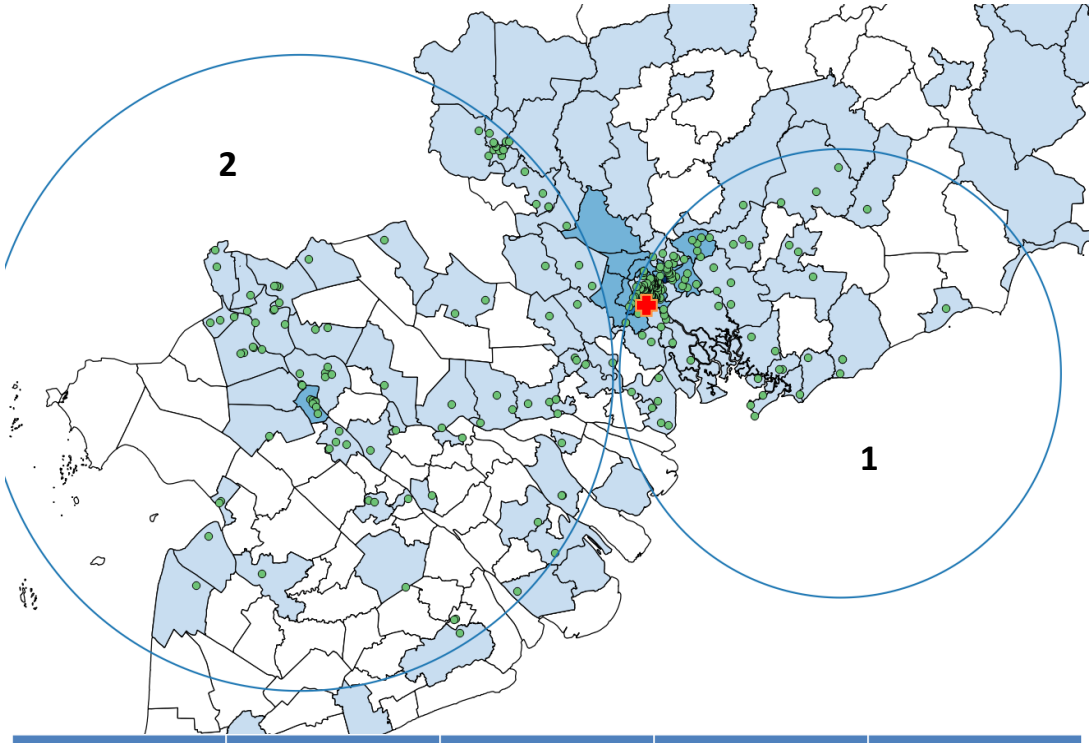


Figure 2-5 Number of cryptococcosis cases in Ho Chi Minh City area at the district level over the period 2000 to 2014. The darker the colour is, the higher the cryptococcal meningitis count in each district. The red cross indicates the site of HTD

2.4.2 Clustering of *C. neoformans* in Vietnam

OverallI identified two clusters of cryptococcal meningitis (Figure 2-6). Of the two clusters, cluster 1 is the most likely cluster – twice as many number of cases as the secondary cluster (cluster 2). Cluster 1 is smaller (radius= 90 km) spanning half of Ho Chi Minh city, Dong Nai and Vung Tau while cluster 2 (radius=127 km) covers most of the Mekong delta. Cluster 1 contained 115 cases occurring over the years 2004 to

2007; cluster 2 contained 61 cases and occurred during the years 2004 to 2011. Within cluster 2, Long Xuyen had the highest burden of cryptococcosis (7 cases).



Cluster	Case	Time frame	Relative risk	P value
1	115	2004 - 2007	2.1	< 0.001
2	61	2004 - 2010	2.25	0.001

Figure 2-6 Temporal–spatial clustering of *C. neoformans* in Vietnam. The radius of cluster 1 is much less smaller than that of cluster 2 (90 vs 127 km). Each dot represents a single isolate from a patient. The red cross indicates the site of HTD

2.4.3 Clustering of *C. neoformans* in Vietnam adjusted by age and gender

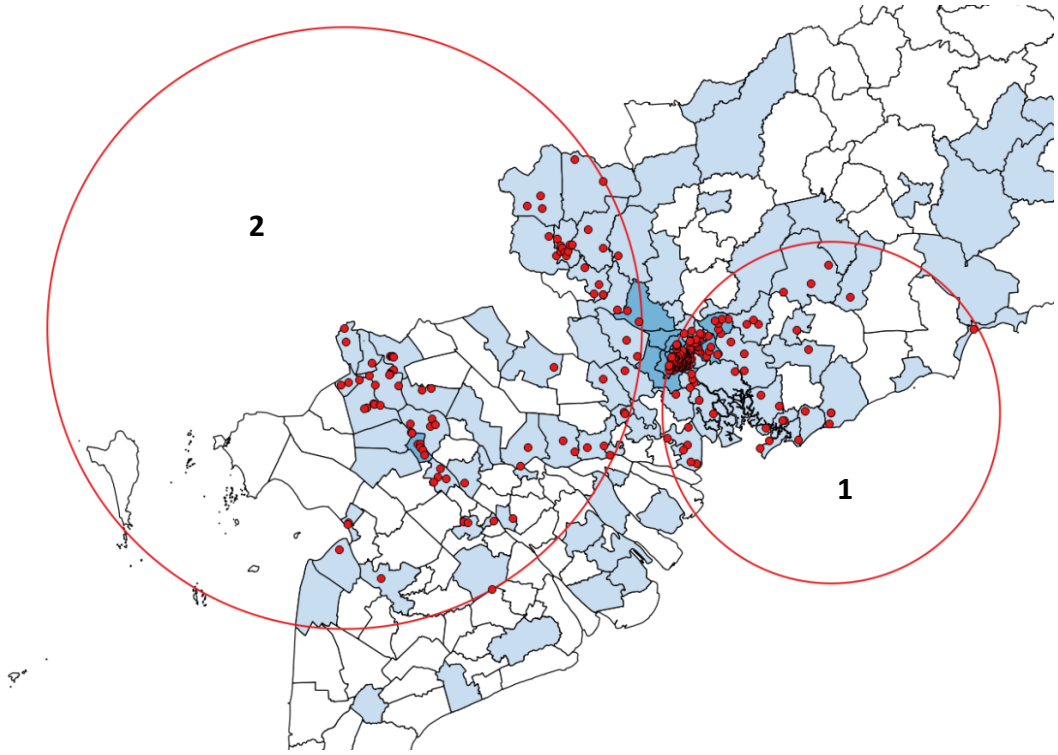
Figure 6 shows temporal-spatial clustering of C. neoformans in Vietnam after adjustment for age and gender. The number of cases in the most likely cluster (cluster 1) remained unchanged. The number of cases in the secondary cluster (cluster 2) drops by 234 % because the time frame shrinks and spans only three years (2004-2006).

Cluster	Case	Time frame	Relative risk	P value
1	115	2004-2007	1.96	0.0002
2	26	2004-2006	3.23	0.01

Figure 2-7 Temporal–spatial clustering of *C. neoformans* in Vietnam. Each dot represents a single isolate from a patient.

2.4.4 Cluster of cryptococcosis by HIV status

Analysis of cases of cryptococcal meningitis associated with HIV revealed that there appeared to be two clusters (Figure 2-8). Cluster 1, the most likely cluster, contained more than four times as many cases as cluster 2. The time frames of these clusters coincided with the peaks of HIV associated cryptococcosis (Figure 2-2)



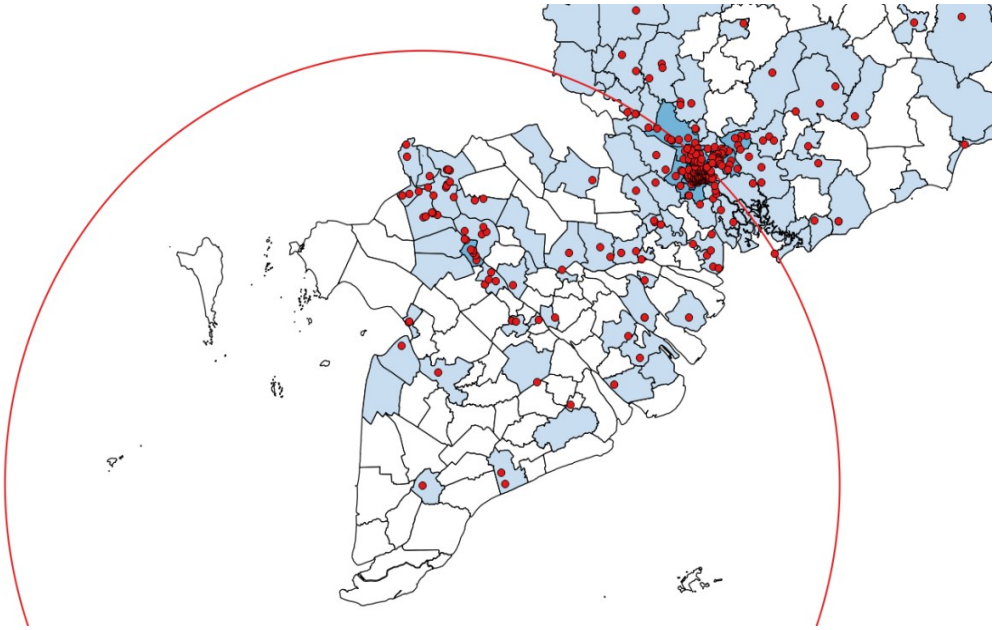
Cluster	Case	Time frame	Relative risk	P value
1	144	2004-2010	2.67	<0.001
2	35	2012-2014	4.25	<0.001

Figure 2-8 Temporal-spatial clustering of *C. neoformans* in HIV patients. The radius of cluster 1 is 1.8 times smaller than that of cluster 2 (88 vs 155 km) but it is more populous than cluster 2 (144 vs 35 cases).

2.4.5 Cluster of cryptococcosis by HIV status adjusted by age and gender

When I adjusted for age and gender, I identified only a single cluster of *C. neoformans* in HIV patients (Figure 2-9). This was large (r=247 km) and spanned more than half of Ho Chi Minh city and the whole Mekong delta. However, the

number of cases in the cluster was smaller than the result before adjustment (87 vs 144).



Case	Time frame	Relative risk	P value
87	2004-2006	2.04	0.0007

Figure 2-9 Temporal-Spatial clustering of *C. neoformans* in HIV patients after adjustment for age and gender.

2.4.6 Cluster of cryptococcosis from HIV-uninfected patient

I also analysed clustering of 59 cases from HIV-uninfected patients from 2000 to 2014. There was no clear temporal-spatial clustering, with a distribution that appeared to be random (P=0.22, before adjustment and P=0.155, after adjustment of age and gender). VNIa-5 was the dominant lineage in HIV-uninfected patients, accounting for 71% (42/59) of all lineages. Additionally, VNIa-5 strains also scatter more widely than other strains (Figure 2-10).

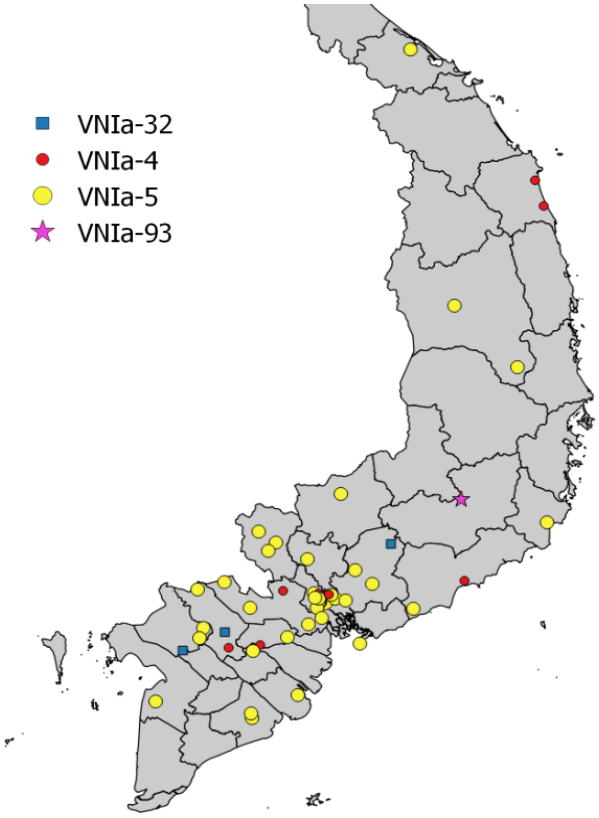


Figure 2-10 Spatial distribution of lineages in HIV-uninfected patients. Dispersal of VNIa-5 isolates is wider than other lineages.

2.4.7 Clustering of cryptococcosis by lineage

C. neoformans was distributed unevenly in Vietnam by lineage. Table 2-1 lists the number of strains of each of the major WGS -derived lineages circulating in our patients. VNIa-outlier and VNIb were seen only in patients from Ho Chi Minh city whereas infection due to VNIa-X strain was only seen in patients from the Mekong delta (Figure 2-11). Disease due to the lineages VNIa-32 and VNIa-93 were more prevalent in Ho Chi Minh city (11/19 and 25/44 cases, respectively) than other provinces (Figure 2-12). Lineages VNIa-4, VNIa-5 and VNIa-32 formed 3 distinct clusters. The lineage VNIa-4 cluster was the largest in terms of number of cases (Figure 2-13), and was to the west of HCMC near the border with Cambodia. Despite the higher frequency of disease due to lineage VNIa-93 (44 cases compared with 19 cases of lineage VNIa-32), I could not detect any evidence of clustering (P = 0.4).

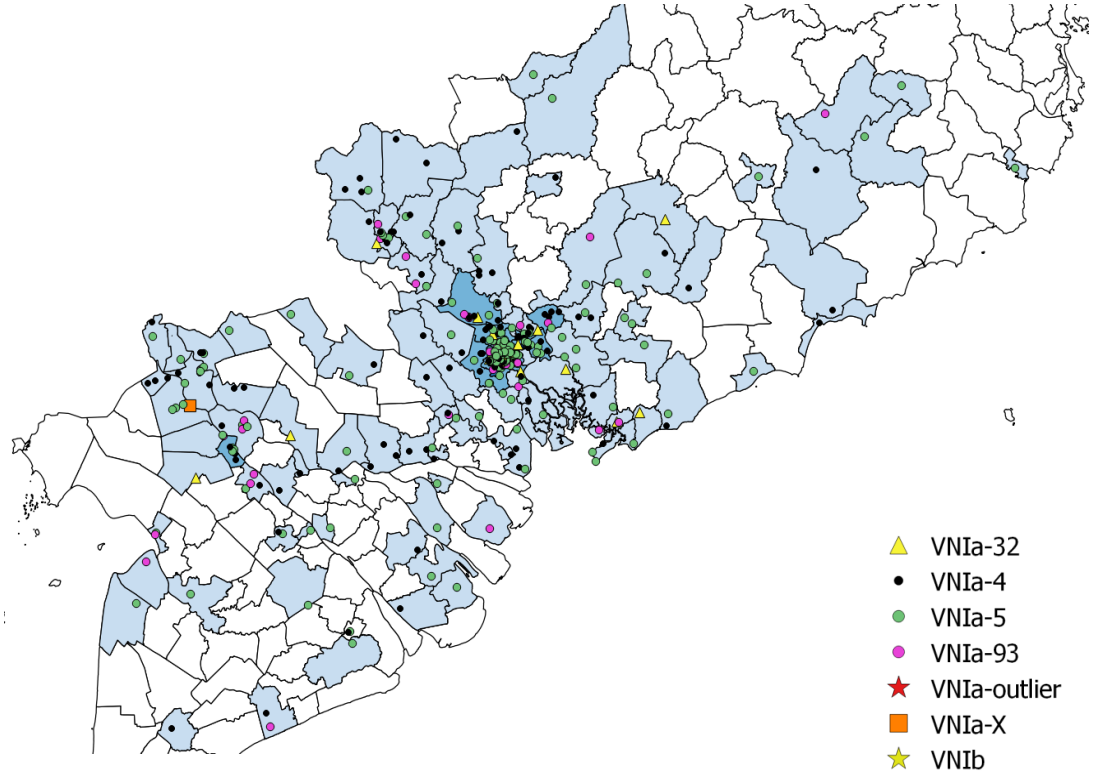


Figure 2-11 Geographical distribution of *C. neoformans* by lineage from 2000 to 2014. Each dot indicates a single strain from a patient.

The lineage VNIa-5 cluster partly overlapped both the lineage 4 and lineage VNIa-32 clusters (Figure 2-13). HIV-uninfected cases were only found in the lineage VNIa-5 and lineage VNIa-32 clusters. Of the three clusters, the lineage VNIa-32 cluster was formed in the shortest time frame (2004-2006).

Lineages VNIa-4, VNIa-5 and VNIa-32 were the most dominant in 2005, accounting for 45% (30/67), 40% (27/67), and 10% (7/67) of all strains in that year (Figure 2-14). Overall, lineage VNIa-4 was the most abundant (179 cases), but lineage VNIa-5 was the most persistent being isolated every year for 15 years (Figure 2-14). In contrast, strains of the lineage VNIa-X, VNIb and VNIa-outlier were scarce with only one case detected and thus it was not possible to demonstrate clustering. Of the three main clusters, the lineage VNIa-4 cluster was the largest (125 km in radius compared to 104 km for the lineage VNIa-5 cluster and 76 km for the lineage VNIa-32 cluster).

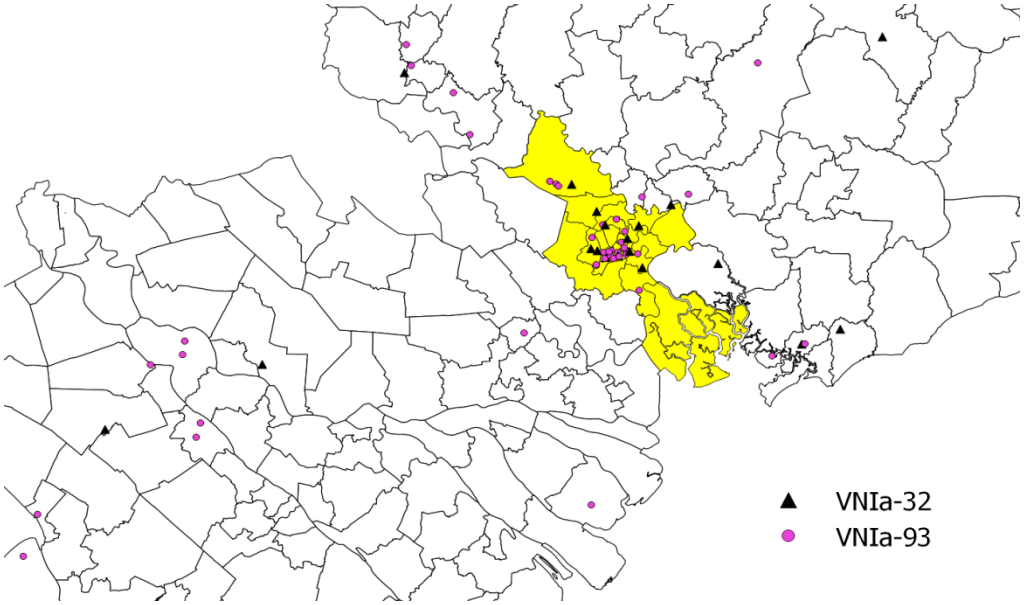
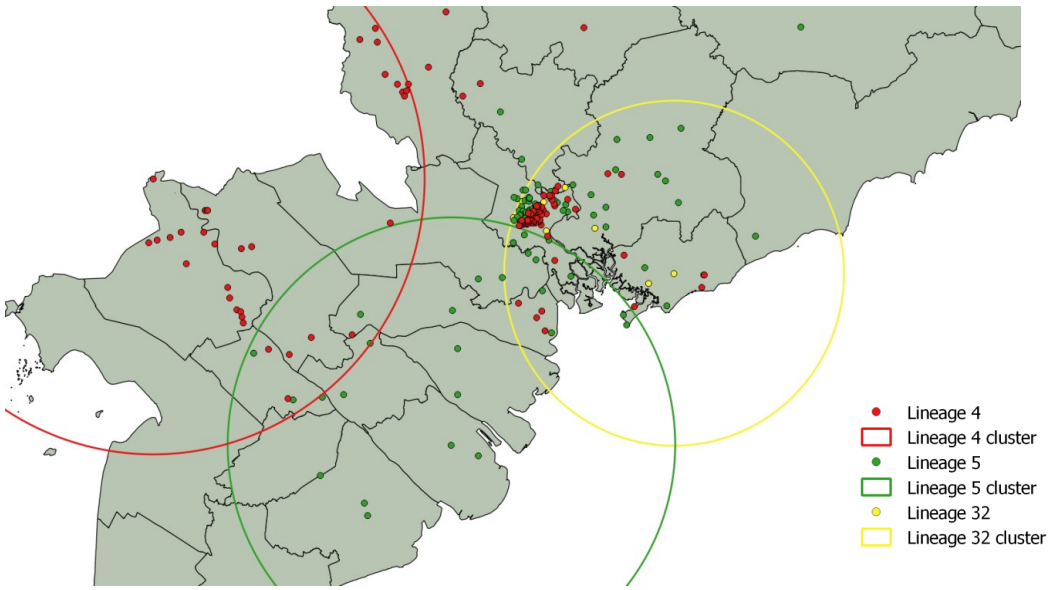


Figure 2-12 Spatial distribution of VNIa-32 and VNIa-93 strains from 2000-2014. The majority of these strains were derived from patients in Ho Chi Minh city



Cluster	Case	Time frame	Relative risk	P value
4	23	2008-2014	3.8	0.002
5	11	2004-2007	10.1	0.0003
32	10	2004-2006	8.7	0.0007
93	7	2004-2009	3.7	0.4

Figure 2-13 Space-time clustering of *C. neoformans* strains in lineage VNIa-4, VNIa-5 and VNIa-32 from 2000 to 2014.

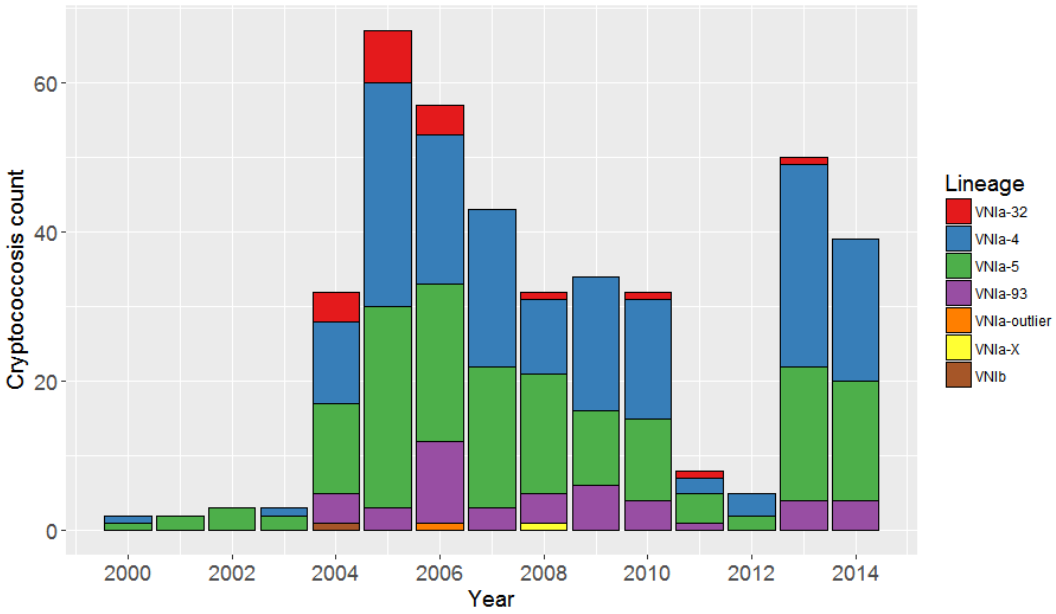
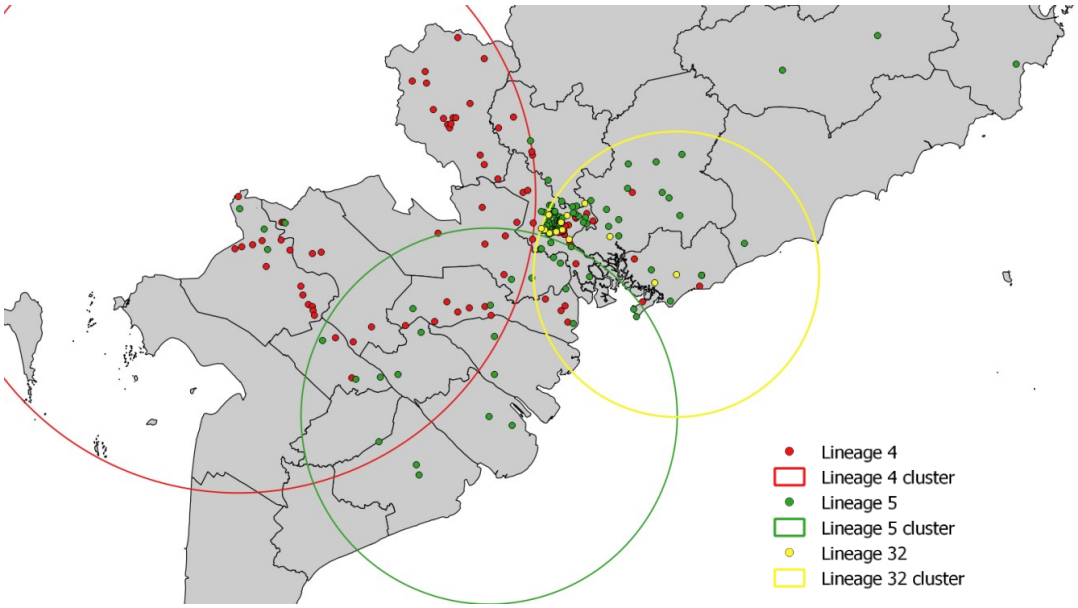


Figure 2-14 Annual distribution of *C. neoformans* by lineage from 2000 to 2014 in Vietnam. Strains of the lineage VNI-5 occurred every year, whereas strains of the lineage VNIb, VNIa-outier and VNIa-X were quite scarce with occurrence in one year for each case.

2.4.8 Clustering of cryptococcosis by lineage after adjustment for age and gender

Figure 2-15 shows temporal-spatial clusters of three lineages (4, 5 and 32) following adjustment for age and gender. The number of cases increases within the lineage 4 cluster after adjustment although the time frame shrinks. The lineage 4 and 32 clusters remained spatially separate but they still overlapped with the lineage 5 cluster. The lineage 4 cluster remained the largest in terms of both cases and radius. Clustering of lineage 93 was not detected.



Cluster	Case	Time frame	Relative risk	P value
4	40	2009-2014	2.8	0.001
5	11	2004-2007	9	0.001
32	12	2004-2006	6.3	0.007
93	3	2005-2006	1.6	1

Figure 2-15 Space-time clustering of *C. neoformans* strains in lineages VNIa-4, VNIa-5 and VNIa-32 after adjustment of age and gender from 2000 to 2014

2.4.9 Clustering of sequence type 5 strains

Strains of sequence type 5 (ST5) are of particular interest because they are associated with HIV-uninfected patients[88]. Cluster analysis of ST5 strains derived from HIV-negative patients (40 cases) was performed but gave no significant cluster, probably because of the wide dispersal and relatively low numbers of HIV-uninfected patients with cryptococcal meningitis (data not shown). Therefore, I analysed the cohort of HIV infected patients for this lineage (N=103). I detected a cluster of 25 cases (relative risk 2.88, P-value=0.017), with a radius of 111km, spanning the southwest of Ho Chi Minh City and the adjoining part of the Mekong delta (Figure 2-16). Within the cluster, 18 out of 25 cases (72 %) came from urbanised areas (Ho Chi Minh City and Vung Tau city), and the remaining 7 (32 %) from predominantly rural environment in the Mekong delta (P = 0.0046). The consistency of this cluster

suggests the prevalence of environmental *C. neoformans* lineage VNIa-5 is significantly higher in urban rather than rural areas. I use the identification of this VNIa-5 cluster for targeted environmental sampling in the next chapter.

Figure 2-16 Space-time clustering of ST5 isolates from HIV-positive patients (25 cases, time frame 2005-2006, P-value = 0.017). The inset features southern Vietnam and part of Cambodia.

2.5 Discussion

I am interested in understanding the epidemiology and ecology of *Cryptococcus neoformans* in Vietnam. As an initial step, I used data from our clinical studies previously performed at the Hospital for Tropical Diseases and Cho Ray hospital, to understand geographical and time trends in this disease in Vietnam. Mapping clinical data can at least give an indication of the distribution of lineages of medical importance. Of note clinical trials at OUCRU are designed within a pragmatic framework, the aim being to ensure that the results of trial findings are relevant to the general patient population. To achieve this, it is essential that the trial population is representative of the wider patient population. This is achieved by minimizing inclusion and exclusion criteria. Because of this, it is likely that the isolate populations used in my studies are essentially representative of the wider population, with the exception of the break in study activity in 2011-2012. People

are infected with C. *neoformans* by inhaling desiccated yeasts or basidiospores from their surroundings, mostly their residence, from early childhood [70]**.** Thus the interpretation of mapping data from clinical cases is complicated by the fact that the incubation/latent period of infection with *C. neoformans* can be long and can be variable. It may even be variable between lineages, particularly if there is variability in the pathogenicity of lineages. I assumed that the majority of patients got infected with *C. neoformans* near their home, and that the majority of patients within my study set continue to live near the site of infection. However, I cannot be sure that this is the case since these data (migration history) were not gathered within the clinical trials, this not being their primary purpose. A further weakness of my study is that there was a break in the recruitment of HIV infected patients during 2011 and 2012 because there was no active clinical trial at that timepoint (HIV uninfected patients continued to be enrolled into a prospective descriptive study of CNS infections). A further potential weakness of my study is that I have assumed that the patient referral patterns to HTD have not changed over time. I had to do this, because it was not possible to obtain hospital referral data by patient address over the 15 years of this retrospective study. However, I have no particular reason to suspect that referral patterns for this disease will have changed much over time. Despite these limitations, an ecological niche of *C. gattii* in British Columbia, Canada was successfully modeled using clinical records followed by validation using nontravel cases (animal and environmental samples)[170] .

2.5.1 Overall distribution of cryptococcosis in southern Vietnam

The prevalence maps (Figure 2-3 and 2-4) show that half of the country was in the catchment area for our hospital (from Hue to Ca Mau). No cases in our hospital were from the big cities of the Central Coast, such as Da Nang and Nha Trang. This may reflect the fact that such patients would have received treatment in local hospitals, although complicated and unusual infections are usually referred on through well-trodden and established referral paths. In provincial hospitals, tuberculosis meningitis (TBM) is the third most common CNS infections. Diagnosis of TBM is complicated due to non-specific manifestations and the need for prolonged specialized culture. The low sensitivity, and lack of awareness of cryptococcal disease

could lead to either completely missing, or late diagnosis, of cryptococcal meningitis. However, tuberculosis is common in Vietnam, and TBM is a diagnosis that is readily considered in patients not responding to treatment. There are well established inter-hospital referral patterns in Vietnam, particularly for TB; therefore most cases of cryptococcal meningitis, if not diagnosed, would at least be referred to large tertiary centres such as the Hospital for Tropcial Diseases or Pham Ngoc Thach TB hospital, and the diagnosis of cryptococcal meningitis is likely to be made. Diagnoses made at the TB hospital are transferred to the Hospital for Tropical Diseases. Moreover, a well-resourced prospective surveillance study of CNS infections in district and provincial hsopitals (2007-2010) found only two cases cryptococcal meningitis out of 1241 patients with syndromes consistent with CNS infection[182]. Therefore, for our region, the south of Vietnam, it is possible that my data in fact do cover the vast majority of cases.

I identified a peak number of cryptococcal meningitis cases in 2005. The President's Emergency Plan for AIDS Relief (PEPFAR) was instigated in Vietnam around 2005 and has had a dramatic effect in delivering antiretroviral therapy (ART) to patients, and subsequently reducing the risk of opportunistic infections. For this to be the reason that the number of cases of cryptococcal meningitis fell in subsequent years implies that the incubation period is short, in the order of a matter of months, for a significant number of patients. Alternatively it may represent a change in referral pattern or health seeking behavior by HIV infected patients around this time, although I have no data to support this occurring.

Most cases of cryptococcal meningitis occurred in Ho Chi Minh city, the Mekong Delta and the Southeast, which are the regions with the most patients living with HIV [183]. Ho Chi Minh has an HIV prevalence of 1.25% as opposed to 0.5% for the country overall [183]. Moreover, HCMC population density is highest in Viet Nam (General Statistics Office; http://www.gso.gov.vn). Taken together, these factors drive the highest burden of cryptococcosis and the most likely clusters of *C. neoformans* spanning HCMC and bordering well-developed provinces in the Southeast - Ba Ria Vung Tau and Dong Nai (Figure 2-7). In the Mekong delta, An Giang, the most populous province, has the highest burden of the disease (Figure 2-6). In southern

Vietnam, the dry season mostly occurs from May to November wherease the monsoon lasts from December to April. There is no evidence about dramatic change of climate during the study period, so climate may not change the overall distribution pattern of cryptococcal meningitis.

2.5.2 Clustering of cryptococcal meningitis in southern Vietnam

I hypothesized that cases of cryptococcal meningitis were randomly distributed in southern Vietnam. However I found two temporal-spatial clusters of cryptococcal meningitis. The most likely cluster (cluster 1) covers half of Ho Chi Minh city and the Southeast (Figure 2-7). After adjustment for age and gender, the size of cluster 1 did not change, suggesting age and gender did not interfere with overall clustering. Cluster 1 may be driven by the high prevalence of HIV in HCMC, where the percentage of adults (age 15-49) living with HIV is the highest in the country (1.25 % compared with 0.51 % for Vietnam as a whole in 2005)[183]

2.5.3 Clustering of *C. neoformans* is associated with HIV

I identified temporal-spatial clustering of cryptococcal meningitis within HIV patients. However, the time frame of the cluster may be driven by the recruitment of cryptococcosis patients during the trial. The cluster covers half of Ho Chi Minh city and all of the Mekong delta. However, age and gender did appear to interact with this clustering (Figure 2-9). HIV-unassociated cases of cryptococcal meningitis appeared to be scattered randomly and more widely than HIV patients (Figure 2-3). HIV-uninfected patients were highly associated with lineage VNIa-5 infections (Table 1). This association suggests that lineage VNIa-5 may have increased ability to cause disease (pathogenicity). Increased pathogenicity may be due to adaptation to an ecological niche. My observation suggests that *C. neoformans* VNIa-5 strains are better able to adapt to different ecological habitat to expand their distribution (Figure 2-3 & 2-10). Expansion of ecological habitats has been occuring with *C. gattii* VGIIa, which caused the cryptococcosis outbreak in Vancouver Islands[10].

2.5.4 Lineage-specific clustering of *C. neoformans*

I found three distinct clusters of VNIa-4, VNIa-5 and VNIa-93, suggesting 3 distinct reservoirs of these lineages in the environment. Of three lineages, the lineage VNIa-32 cluster was formed in the shortest time frame, suggesting the outbreak of *C. neoformans* lineage VNIa-32 strains within this time (2004-2006).

Lineage VNIa-93 strains were more prevalent than lineage VNIa-32, however only lineage VNIa-32 strains significantly cluster. The random distribution of lineage VNIa-93 may indicate their wide dispersal in the environment. Lineage VNIa-93 is a highly clinically significant cause of disease in Africa, where it is the most frequently isolated strains with 61% of isolates belonging to this lineage [184], but it is a relatively infrequent cause of disease in Vietnam (Ashton *et al,* manuscript under review). The lack of clustering seen with this strain would be consistent with imported disease; however, I lack travel histories for my patients although frequent travel to and from Africa by Vietnamese during the time period of this study was uncommon. Instead, these strains could be dispersed into Vietnam either by birds or the wind. Infection with lineage VNIa-93 is associated with a significantly reduced risk of mortality by 10 weeks (Hazard Ratio = 0.45, p = 0.003) (Ashton *et al,* manuscript under review). The clustering of lineage VNIa-32 around Ho Chi Minh city suggesting these isolates are adapted to urban area (Figure 2-13). The cluster of lineage VNIa-4 disease that I identified in the Mekong Delta consisted of a large number of strains over an extended period (6 years after adjustment of age and gender). This suggests the Mekong delta is a favoured ecological niche for this lineage. This cluster extends to the Cambodian border; in Thailand MLST types ST4 and ST6 account for the vast majority of infections (Thanh *et al*, manuscript under review), and thus clustering to the west of Ho Chi Minh City is perhaps not surprising. The lineage VNIa-4 cluster had the largest area (radius of 164 km), implying a widespread favourable habitat for this isolates disperse. The persistence of lineage VNIa-5 infection over 15 years, with consistent recovery of disease due to this lineage, implies it is well adapted and indigenous to the Vietnamese environment.

Spatial clustering detection has been used to investigate the epidemiology of dengue, malaria, encephalitis, typhoid fever, leukemia incidence, etc. [177,185–187] .This

is the first study to investigate temporal-spatial clustering of cryptococcal meningitis caused by *C. neoformans* (VNI). To the best of my knowledge, my study has the largest sample size of any cluster detection study (409 isolates derived from 409 locations). This has been enabled by the large number of cases of disease enrolled into clinical trials in a single center. Litvintseva *et al* found *C. neoformans* (VNB) is endemic to southern Africa and associated with the mopane tree (*Colophospermum mopane*)[56]. Previously, the ecological niche of the sibling species *C. gattii* responsible for the outbreak in Vancouver Island, British Columbia, Canada in 1999 was predicted using human and animal cases and positive environmental locations in the endemic area[170]. *C. gattii* isolates clustered along the central and south-eastern coast of Vancouver Islands. They generally colonized acidic soil with low moisture and low organic carbon contents [53].

The method I used has limitations in that it is unable to detect irregular shaped clustering due to its circular scanning approach [188]. The circular scanning tends to identify a clusters that are larger than the true cluster by including surrounding areas without elevated risks [189]. Clustering may be confounded by the fact that I assumed people working in Ho Chi Minh city reported their permanent addresses in their hometown or elsewhere, and that I did not have travel or immigration histories for patients. The patients' residence may not be the place of infection, and there may be long latency between infection and clinical presentation[72], making the identification of cluster uncertain. When inhaled in lung, the yeasts are likely to be in dormant forms for a long time before they are reactivated upon host immune deficit[72]. In the sibling species *C. gattii*, the duration between exposure to the pathogen and onset of cryptococcosis symptoms ranges from 2 to 11 months[69].

2.6 Conclusion

I found uneven distribution of cryptococcal meningitis in Vietnam, suggesting that some areas are relative hot spots for disease and that there is niche adaptation by the organism. Ho Chi Minh city has the most cases. I also detected wide dispersal of lineage VNIa-5 strains, which implies either that their favourable niche is widespread, or that this lineage can adapt to multiple different environments. This increased adaptability may also confer enhanced pathogenicity, allowing adaptation

to the immunocompetent host. I could not detect evidence of clustering of lineage VNIa-5 strains in HIV uninfected patients, although clustering was detected when I looked at strains associated with HIV infected patients. The failure to show clustering in HIV uninfected patients may have been due to lack of power because of the relatively low number of cases. Three distinct clusters of lineage VNIa-4, lineage VNIa-5 and lineage VNIa-32 were identified, suggesting their adaptation to different ecological habitats.

Chapter 3

Environmental sampling and characterisation of environmental *Cryptococcus spp*

3.1 Introduction

Cryptococcus neoformans is an environmental saprophyte, and acquisition of infection is believed to be the result of inhalation of an infectious propagule (spore or yeast) directly from the environment. The lack of person to person spread means that it is unlikely that the ability to infect humans has been driven from within the human host. Rather, the pathogenicity of *Cryptococcus* has been considered to be a 'bystander' effect of its adaptation to its ecological niche. The emergence of different lineages likely reflects further adaptation to specific environmental conditions. Variability in pathogenicity between lineages, such as that exhibited by ST5 (i.e. causing disease in apparently immunocompetent hosts) is presumably a consequence of such specific niche adaptation.

The ecology of *Cryptococcus* species is incompletely described. Established ecological habitats include avian droppings (particularly, chicken and pigeon), soil, and decayed woods of various tree species, although these have not been interrogated to the resolution of MLST. Concordance has been established between clinical and environmental isolates recovered from the same geographical areas. For example, clinical and environmental isolates of *C. neoformans* var. *neoformans* had identical RAPD profiles in Nagasaki [190]. RAPD and PCR fingerprinting confirmed *C. gattii* (VGI) strains were the molecular group predominantly isolated in eucalypts (100%) and in clinical settings (92%) in Australia [191]. In southern Africa, endemic *C. neoformans* var. *grubii* isolates (VNI and VNB) are associated with mopane trees (*Colophospermum mopane*) with 30 % of sampled mopane trees colonized by *C. neoformans* var. *grubii* [56].

Some studies have found that the environmental recovery rate of isolates varies by season, (dry versus wet), and varies from region to region[26,57,192]. It is frequently high in regions with high prevalence of cryptococcosis but low in others. For example, the

recovery rate is up to 32 % in southern Africa but only 0.15 % in Havana, Cuba[193] [56]. In northern Thailand, it has been reported that *C. neoformans* is more frequently isolated from the environment in the dry season compared with the rainy season[57]. However, this contrasts with the situation in Bogota, where the rainy months favour the successful isolation of *C. neoformans* from the environment.[194] Of note, there does not appear to be seasonal variability in the incidence of human disease, likely representing a smoothing effect of latent infection [195].

It is not clear whether identifying the ecological niche of *C. neoformans*, or particular lineages, would have a pragmatic implication for public health. However, understanding the niche may provide insight into the drivers of pathogenicity. Of note, previous studies using environmental sampling may have been biased because sampling was not randomized [56] [171] [26]. Randomised sampling protects against confirmation bias when attempting to define the ecological habitat of C. *neoformans*, and should identify the actual prevalence and diversity of different genotypes in the whole population. A potential disadvantage of randomized sampling is that it is resource intensive and may produce a lower yield of isolates. In this chapter, I aim to understand the ecology of *C. neoformans* in Vietnam, identify whether clustering occurs by species or lineage, and identify specific environmental niches that will provide insight into the origins of bystander pathogenesis. I use both random and targeted sampling, in both urban and rural areas, in an attempt to provide comprehensive and unbiased sample datasets.

3.2 Aims

My aims were:

1. To describe the diversity of *Cryptococcus spp.* in the environment
2. To determine whether geographical clustering of *C. neoformans* isolates occurs by MLST defined lineage
3. To determine whether any association occurs between *Cryptococcus neoformans* lineages and tree species
4. To characterize the virulence-associated phenotypes of environmental-sourced *Cryptococcus* species

3.3 Materials and Methods

3.3.1 Randomised sampling

I collected environmental samples consisting of tree swabs, sub-canopy soil and air. I chose two main types of environment –urban and rural (consisting of mixed farming including rice, rubber and village environments). Each sampling area was viewed in quantum GIS with a Google Earth plug-in (www.qgis.org) and overlayed with a 22 by 22 km grid (Ho Chi Minh) or 13 by 11 km grid (Nhon Trach) bisected by gridlines 100 metres apart using quantum GIS 1.7.4. For each grid I randomly selected 500 gridline intersects using R 3.0.0 (Figure 3-1 & 3-2). The chosen intersects were marked as waypoints in a Garmin Montana 600 Global Positioning System (GPS) (Garmin Ltd, Kansas, USA) to enable navigation to the precise sampling site. At each site, following explanation of the project and obtaining the verbal consent of the landowner, the nearest tree to the selected coordinate was sampled. I collected two tree swabs, two soil samples and one air sample at each site (1000 sites altogether). Tree species were identified using Northern Annamites Tree V.1.0 (National University of Laos). Nhon Trach was sampled in the early dry season to monsoon (April to August 2014) while Ho Chi Minh City was sampled from late monsoon to dry season (November 2014 to January 2015).

Sterile transport swabs containing Amies medium without charcoal were used to swab tree hollows, and overlying bark. A single swab was wiped repeatedly until darkened on all sides. Air was collected under each tree canopy for 10 minutes using an Oxoid air sampler (Oxoid, Hampshire, UK) with a flow rate of 100 litres of air per minute (1000 litres in total). Soil samples was collected within the tree root zone, around 5 cm deep from the surface and contained in clean ziplock bags.

Samples (soil and swab) were placed initially in a cool box for transportation and then stored at air conditioned room temperature (25^{o}C) and processed as soon as possible but no later than one week following collection. At each sampling site, the global positioning system receiver (Garmin) was used to record the longitude and latitude of the sampling site.

3.3.2 Targeted sampling

Because of the potential risk of low yield through randomized sampling, I also undertook targeted sampling. Targeted sampling, while not likely to deliver a set of isolates that are representative of the entire population of *Cryptococcus* in Vietnam, has the advantage of potential higher yield. Targeted sampling had 2 themes – sampling environments that I considered were potential habitats based on my knowledge of the biology of *Cryptococcus spp*, and sampling based upon my previously identified clusters of clinical cases (chapter 2). Targeted samples included tree swabs, soil and air, as described above. Each site had two tree swabs, two soil samples and one air sample collected.

Biological sampling I sampled gecko (n=30) and avian droppings (N = 200; pigeon, chicken, wild bird, and duck), and fruits (n=25). Because *Cryptococcus* infection is thought to be due to the inhalation of airborne infectious propagules, I sampled the swabs of fan blades and air conditioner filters, assuming that this machinery processed large volumes of air and might therefore serve as a concentration point for trapping airborne yeast particles or spores (n=30). In addition I sampled a more speculative source, potentially relevant in Vietnam - rice husk (n=20).

Clinical cluster sampling I also selected sampling sites based on the result of the cluster analysis of the 103 clinical isolates (ST5) derived from HIV patients (Chapter 2). I went to 25 of the clinical-defined cluster sites and collected environmental samples as described above.

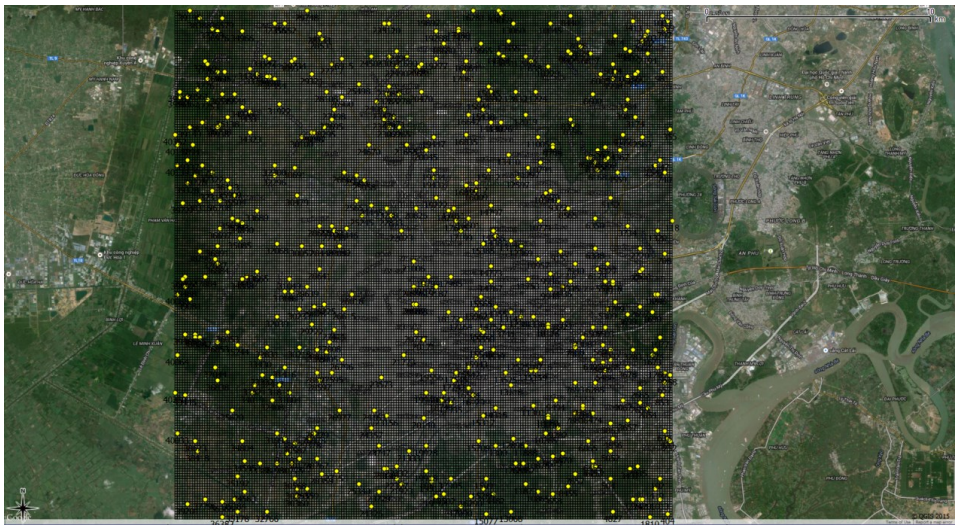


Figure 3-1 A gridline (22 x 22 km) overlaying Ho Chi Minh city satellite imagery with 500 randomised sampling spots (yellow dots) visualised by QGIS 1.7.4

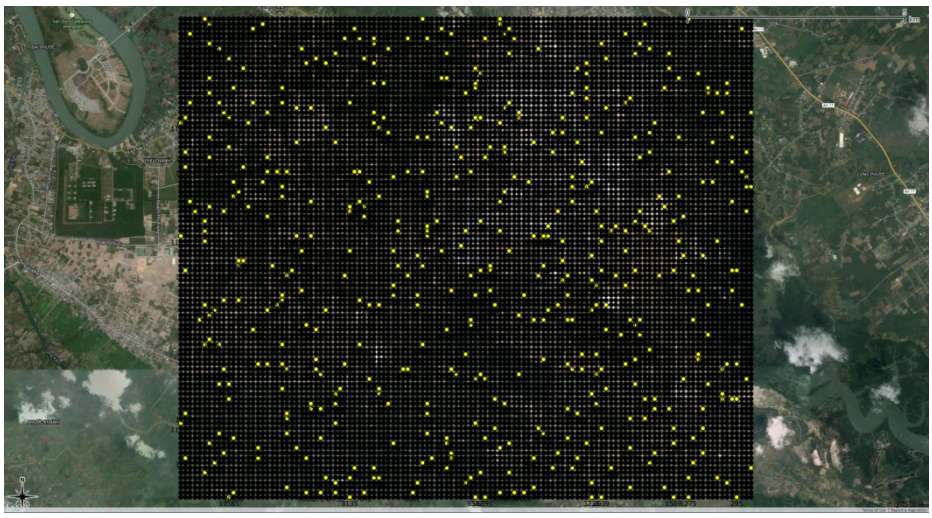


Figure 3-2 A gridline (13x11 km) overlaying Nhon Trach with 500 randomised sampling spots (yellow dots) visualised by QGIS 1.7.4

3.3.3 Isolation

Soil and avian/gecko excreta

Soil was collected within a tree's root zone (tree canopy) at 5-10 cm depth. Samples were processed as described by Litvintseva *et al* [171]. I suspended 2 - 3 grams of soil or guano in 15-mL sterile water by vortexing at 2500 rpm for 2 minutes. The sediment was settled for 5 minutes, and then 100-µL aliquot of the supernatant was plated onto a Benomyl Bird seed agar plate. Benomyl was added to inhibit ascomycetes growth (1.5 ug/mL). It has no effect on basidiomycetes[196]. I decant 8 mL of

supernatant into another 15-mL Falcon tube and centrifuge at 3220 G in 15 mins. I poured away 6 mL of the supernatant, resuspended the pellet by vortexing at 2500 rpm for 2 minutes, and inoculated 100-µL of the suspension onto another modified Bird seed agar plate. Incubate all plates at 30^{o}C for at least 3 days. I observed for the appearance of characteristic brown colonies from 3 days until 7 days after inoculation.

Bark and tree hollow swabs

Collected swabs were thoroughly agitated in 6-mL sterile water and then vortexed at 2500 rpm for 2 minutes. 100-µL aliquot of the resulting suspension was inoculated onto a benomyl bird seed agar plate. I then centrifuged the remaining suspension at 3220 G for 15 mins, poured off 4-mL of the supernatant, resuspended the pellet and inoculated 100 µL of the new suspension onto another benomyl bird seed agar plate. All plates were incubated at 30^{o}C for at 7 days. I observed for the appearance of mucoid or characteristic brown colonies from 3 days until 7 days after inoculation.

3.3.4 Bird seed agar

Growth on bird seed agar results in the development of brown colonies due to the production of melanin. This medium permits selective recovery of *C. neoformans* but inhibits growth of mold and bacteria[197]. All mucoid colonies (white, pale or dark) were examined for identification of presumptive *Cryptococcus species (oval to round encapsulated budding yeast forms).*

3.3.5 Biochemical speciation

The ability of *Cryptococcus* species to assimilate carbohydrate substrate was used for identification using ID 32 C strips (bioMerieux, France). Kits were used according to the manufacturer's instructions. Briefly, a single colony of *Cryptococcus spp* grown in Sabouraud agar was suspended in sterile water to make a suspension with a turbidity equivalent to 2 McFarland (compared with a turbidity of 2 McFarland standard). Dispensing 135 µL of the suspension into each well of the strips followed by incubation at 30^{o}C for 72 hours and then growth is detected by visual reading. Fungal identification is obtained using the Api-web identification software https://apiweb.biomerieux.com

3.3.6 Sequence-based speciation

3.3.6.1 DNA isolation

Isolates were revived on Sabouraud's agar at 30^{o}C for 72 h. Single colonies were spread for confluent growth and again incubated at 30^{o}C for 24 h. Chromosomal DNA was extracted in accordance with the previously described protocol by Wen et al. [198]. The resulting DNA pellet was suspended in 100 ul of TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0) containing 10 ug of RNAse A.

3.3.6.2 DNA sequencing

The nuclear eukaryotic rRNA cistron has been used for fungal phylogenetics for more than 20 years[199]. It consists of 18S, 5.8S AND 28S rRNA genes (Figure 3-3). During post transcriptional processing, the cistron splits to remove two internal transcribed spacers. These two spacers and the 5.8S are called the ITS region. This highly variable region is the most frequently used barcode marker of fungi[199]. The internal transcribed spacer (ITS) region and the intervening 5.8S rRNA gene were amplified using universal fungal primers ITS1 (TCCGTAGGTGAACCTGCGG) and ITS4 (TCCTCCGCTTATTGATATGC), as described by White et al[200]. These primer pairs are complementary to the conserved regions at the end of the 18S and 28S rRNA genes, respectively, leading to amplification of the ITS region and the 5.8S rRNA gene. Reactions were performed in a 50 ul reaction mixture containing approximately 100 ng template DNA, 0.2 mM of each primer, 0.2 mM each dATPs, dTTP, dGTP, dCTG; 2.1 mM $MgCl_2$, 1 U of Taq polymerase (Bioline) and 5 µl of 10X buffer. The amplification reactions were performed in a Flexcycle2 (Analytik Jena), with the following cycling parameters: initial denaturation step (94^{o}C for 3 mins), followed by 35 cycles of 94 ^{o}C for 30 s, 55 ^{o}C for 30 s, 72 ^{o}C for 1 min; and a final extension step at 72 ^{o}C for 10 min. PCR amplicons were purified by QIAquick PCR purification (QIAGEN) and sequenced with ITS1 and ITS4 primers using the BigDye terminator v.3.1 reagent kit (Applied Biosystems) in an automated sequencer (ABI PRISM 3130xL Genetic Analyzer, Applied Biosystem).

Figure 3-3 Diagram of the fungal rDNA gene cluster. The barcode region used for fungi identification spans ITS1, ITS2 and 5.8S regions

3.3.7 Cluster analysis

The case-control Bernoulli method[181] integrated in SaTScan v9.4 was used for detecting spatial clustering of *Cryptococcus* species recovered from random sampling. The significance of any detected cluster was assessed by the likelihood ratio test, with a p-value obtained by 999 Monte Carlo simulations generated under a null hypothesis of a random spatial distribution. Detected clusters were considered significant where the resultant p-value was less than or equal to 0.05.

3.3.8 Characterization of virulence-associated phenotypes

Suspected colonies of *Cryptococcus* (beige to brown colonies consisting of oval to spherical yeast like cells) were purified by subculturing onto Sabouraud's agar at 30^{o}C for 2 days and subjected to the following virulence associated phenotypic tests: India ink staining for capsule presence, urease production on Urea Christensen's agar, temperature dependent growth[92], capsule production, melanin production and *ex vivo* CSF survival [201]. The significance of these virulence factors has been described in section 1.5 of the introduction chapter.

3.3.8.1 Control strains

Isolates used in phylogeny and phenotyping experiments are presented below

Isolates	MLST type	Background	Experiment used
H99 ATCC 208821	2	Isolated from male with Hodgkin's disease from USA	Phylogeny of environmental isolates
***Candida parapsilosis* ATCC 22019**	NA	NA	Outgroup in phylogeny, Negative control for urease activity
BK59 **BK48** **BK139**	4 4 5	Isolated from HIV-associated cryptococcosis patients in Vietnam	CSF *ex vivo* survival and Galleria infection model
ΔENA1	2	Derived from H99 strain, growth defect in CSF	Negative control for CSF *ex vivo* survival

Table 3-1 Control isolates used in phylogeny and phenotyping

3.3.8.2 Urease production

Most *Cryptococcus neoformans* are urease positive[135]. Their growth on Christensen's agar results in a characteristic pink colour due to elevated pH driven by degradation of urea by urease. A single colony of suspect *Cryptococcus spp* was streaked onto Christensen's agar followed by incubation from 24 to 48 hours. *Candida parapsilosis* ATCC 22019 was used as negative control.

3.3.8.3 Temperature-dependent growth

Cryptococcus spp were cultured in Sabouraud agar for 2 days at 30^{o}C. Stock suspension of 10^6 cells/mL was made using sterile water. I prepared inoculum by diluting the stock from 10^5 to 10^3 cells/mL. I dropped 5 µL of diluted inoculum to Sabouraud agar and incubated them at 30^{o}C, 37^{o}C and 39 ^{o}C for 72 hours.

3.3.8.4 Capsule production

DMEM agar was used to induce capsule enlargement *in vitro*. Cryptococci inoculated in DMEM plate was incubated at 30^{o}C with supplement of 5 % CO_2 for 5 days. Suspension with Indian ink was made and visualized at 100X magnification using CX41 light microscope (Olympus, Japan). Images were captured using a built-in DP71 camera (Olympus, Japan) and visualised using ImageJ (https://imagej.nih.gov/ij/index.html). Capsule size was calculated by subtracting cell body diameter from the whole cell diameter (including cell body and capsule). 40 individual microscopic cells were measured for each isolate. Two strains of unidentified *Cryptococcus* and three strains of other *Cryptococcus* were selected for capsule enlargement induction.

3.3.8.5 Melanin formulation

L-DOPA agar was used to visualise melanization of cryptococci. I spotted 5 µL of cell suspension (10^6 cells/mL) onto the agar plates followed by incubation for 3 days at 30^{o}C.

3.3.8.6 *Ex vivo* CSF survival

I prepare pooled CSF supernatant by randomly selecting 12 tubes (baseline CSF) prior to filtering. Next I prepared inoculum of 10^7 cells/mL using Cellometer X2 (Nexcelom). 10 uL of inoculum was added to 90 uL of CSF, YPD broth and PBS followed by incubation at 30^{o}C for 3 days. Resulting culture was serially tenfold diluted with PBS (neat, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}). I dispensed dilutions onto Sabouraud plates (5 uL/drop) and exposed for another 3 days at 30^{o}C. Finally I calculated CFU of each strain post-inoculation in CSF. Tested trains were inoculated alongside ΔENA1, negative control (defect in CSF survival)

3.3.9 Sequencing analysis

ContigExpress software was used to obtain consensus sequences from aligned forward and reverse sequence reads. Each resulting sequenced product was submitted to the Basic Local Alignment Search Tool (BLASTN) to retrieve species with identity values > 99 %.

3.3.10 Phylogenetic analysis

Consensus sequences were aligned with the Clustal W2 algorithm [202]. For neighbour-joining analysis [203], the evolutionary distances between sequences were calculated using Kimura's two-parameter model using Bioedit software[203]. Bootstraps analyses were performed with 1000 random resamplings[203] . The tree was rooted using *Candida parapsilosis* ATCC22019 as the outgroup.

3.3.11 MultiLocus sequence typing (MLST)

I determined the MLST genotype of environmental isolates by sequencing 7 unlinked housekeeping genes: CAP59, GPD1, LAC1, PLB1, SOD1, URA5 and GPD1 according to the ISHAM MLST consensus scheme as described by Meyer *et al* [13]. The seven loci were sequenced using ABI 3130xL Genetic Analyzer (Applied Biosystems, USA). Sequence reads (forward and reverse) were assembled using ContigExpress. Allele types of housekeeping genes and corresponding sequence type were assigned by the fungal MLST Database http://mlst.mycologylab.org/

3.3.12 Virulence phenotype in the *G. mellonella* infection model

The pathogenic potential of environmental isolates was assessed using the *G. mellonella* infection model. An inoculum of 10^6 cells (10^8 cells/mL) was injected to last left pro-leg of each larva and incubated at 30°C and 37°C for 10 days. Each isolate was infected with 15 larvae. Further details of the infection model are described in Chapter 4.

3.3.13 Statistical analysis

Statistical analyses were performed using R version 3.1.2 (R Foundation for Statistical Computing, Vienna, Austria) to test equality of proportions with continuity correction. Fisher's exact test was used to determine associations between melanin production and ability to grow at 37°C. 95% confidence intervals for binomial proportion were estimated with Clopper-Pearson method employed to define associations between particular tree species and *Cryptococcus spp*. A P value of less than 0.05 was considered statistically significant. Corrections for multiple comparisons were performed using the Benjamini-Hochberg method. The Cox

proportional-hazards model was employed to compare survival of *G. mellonella infected* with other isolates.

3.3.14 Experimental workflow

Experimental workflow for this chapter is depicted below

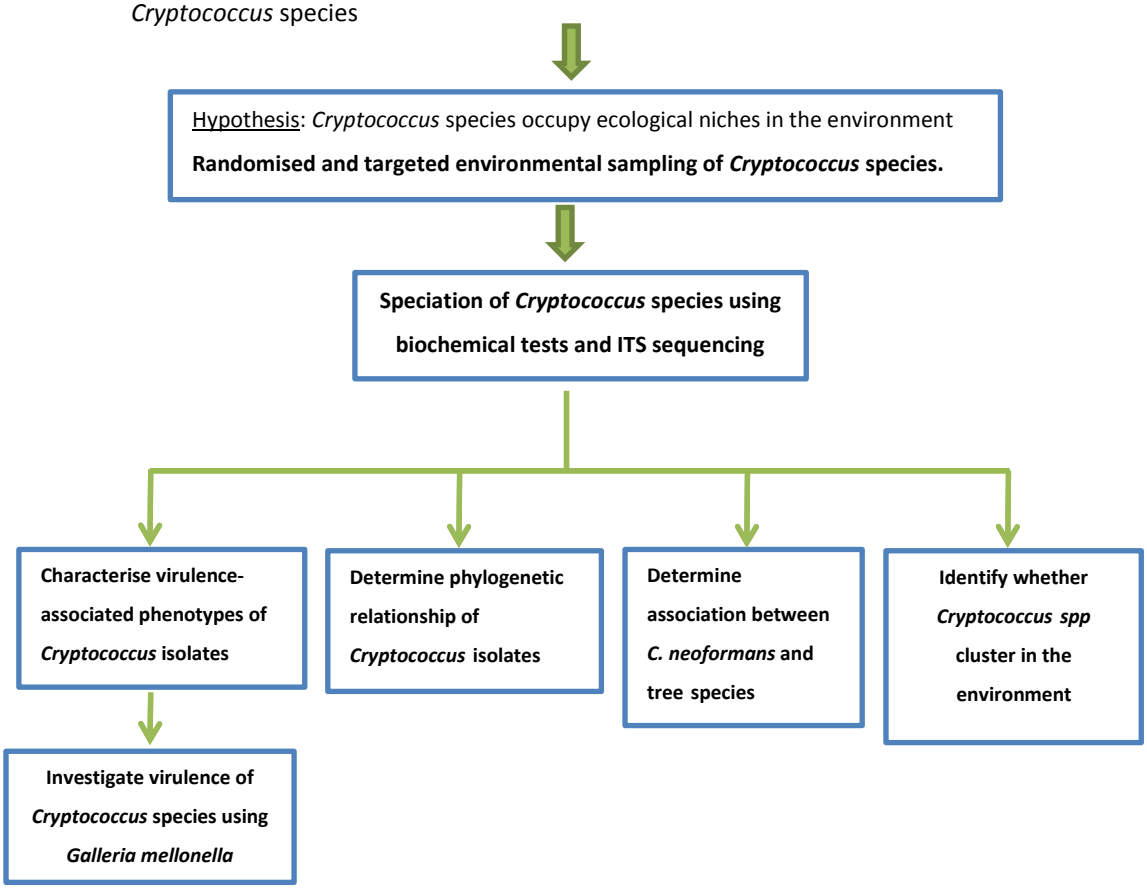


Figure 3-4 Experimental workflow of chapter 3

3.4 Results

3.4.1 Growth on bird seed agar

I identified 123 potential Cryptococcus species from the environmental sampling. Bird seed agar is used for initial differentiation of *C. neoformans* from other *Cryptococcus* species. Thus, of 123 presumptive *Cryptococcus* species, only three *C.*

neoformans exhibit typical brown colour in this agar. The rest were white or pale (Figure 3-5). Under microscopes, non-neoformans cryptococci are oval while *C. neoformans* are spherical yeast-like cells

Figure 3-5 *Cryptococcus* growth in bird seed agar at 30^{O}C. Brown colonies: environmental *C. neoformans*. White or pale colonies: non-*neoformans* species.

3.4.2 Speciation of *Cryptococcus* species

Of the 123 presumptive cryptococci, 107 strains were obtained from random sampling and 16 from targeted sampling (Tables 3-2 & 3-3). All were subjected to carbon assimilation testing (ID32 C, Biomerieux). The speciation of the 123 isolates by ITS sequencing and sugar assimilation tests are shown in table 3-3. There was poor correlation between the 2 methods: only *C. neoformans* strains were concordantly identified by both methods. All non-*neoformans* cryptococci strains were identified as *C. laurentii, C. albidus* and *C. humicolus* by ID32 C (Table 3-4). Using ITS fungal barcoding primers I found that of 123 *Cryptococcus spp*, 3 strains were identified as *C. neoformans* var. *grubii*, 19 *C. laurentii* strains, 93 *C. heveanensis* strains, 6 *C. rajasthanensis* strains and 2 novel *Cryptococcus sp* (they were not previously characterised). Two isolates from Nhon Trach, previously assigned as *C. laurentii* by ID32 C, were later identified as novel *Cryptococcus* species by ITS sequencing. 86 % of strains were misidentified by ID32 C system (Table 3-5). This suggests biochemical testing alone is unreliable for fungal speciation, at least for non-pathogenic species.

Tree name	C. heveanensis	C. laurentii	C. neoformans	C. rajasthanensis	Undetermined Cryptococcus	Total
Acacia auriculiformis	1		1			2
Acacia mangium	**1**					**1**
Alstonia scholaris	1					1
Artocarpus altilis	**1**					**1**
Azadirachta indica	2					2
Cassia fistula	**10**			**1**		**11**
Cinnamomum camphora	1					1
Delonix regia	**18**	**1**				**19**
Dipterocarpus alatus	4	1				5
Eucalyptus globulus	**1**					**1**
Hevea brasiliensis		1				1
Hopea odorata	**6**	**1**	**2**			**9**
Khaya senegalensis	1	2		1		4
Lagerstroemia speciosa	**3**	**1**		**1**		**5**
Manihot esculenta					1	1
Mimosa pigra					**1**	**1**
Mimusops elengi	2	1				3
Muntingia calabura	**2**	**1**				**3**
Peltophorum pterocarpum	11	1				12
Pterocarpus macrocarpus	**3**					**3**
Ravenala madagascariensis	1					1
Samanea saman	**5**			**1**		**6**
Tamarindus indica	2	2				4
Terminalia catappa	**1**					**1**
Tree stump		3				3
Unidentified	**6**					**6**

Table 3-2 107 *Cryptococcus* isolates from random sampling in Nhon Trach and Ho Chi Minh city

Source	Isolate		
	C. heveanensis	C. laurentii	C. rajasthanensis
Artocarpus heterophyllus		**1**	
Bougainvillea spectabilis		1	
Delonix regia	**1**		
Ficus rumphii	1		
Dipterocarpus alatus	**1**		
Hopea odorata			1
Lagerstroemia speciosa	**1**		
Mimusops elengi		1	
Peltophorum pterocarpum	**1**		
Plumeria rubra	1		
Pterocarpus macrocarpus	**1**		
Rhizophora apiculata	1		
Unidentified	**1**		
Air conditioner filter swab			1
Gecko guano	**1**		
Wall swab		1	

Table 3-3 A total of 16 *Cryptococcus* isolates were recovered from targeted sampling sites. The tree species where two isolates were isolated from air are shown in **bold**.

ID32 C speciation result	ITS sequencing speciation result				
	C. neoformans var. grubii	C. laurentii	C. heveanensis	C. rajasthanensis	undetermined Cryptococcus
C. neoformans	3				
C. laurentii		14	88	3	2
C. albidus		4	5	2	
C. humicolus		1		1	

Table 3-4 Correlation between ITS sequencing species identification and ID32 C species identification. Rows represent strains identified by API ID32 C. Columns represent strains identified by ITS barcoding.

Comparison parameter	% (Number/Total number) of strains	
	ID32C	Sequencing
Correct identification	14 (17/123)	100 (123/123)
Misidentified	86 (106/123)	

Table 3-5 Comparison of cryptococcal identification using ID32 C vs ITS sequencing

3.4.3 Distribution of environmental isolates in Nhon Trach and Ho Chi Minh city

Overall, 1000 random sampling sites resulted in the collection of 5000 samples plus 125 samples from 25 targeted sites. I also had 305 samples from gecko guano, avian guano, rice husk, swabs (fan blades, wall, air conditioner filter).

The distribution of environmental *Cryptococcus* species is shown in Figure 3-6. The majority of isolates were *C. heveanensis*; this species is not known to have caused human disease. Of the *C. neoformans* strains, two were isolated from a heavily urban district of Ho Chi Minh city, and one from a rural area of Nhon Trach.

Isolates were derived from both arboreal and non-arboreal sources. Non-arboreal sources included air conditioner filter swab, wall swab, rice husk, fruits and gecko guano. This finding implies cryptococci are ubiquitous in the environment. For random sampling, sites from Ho Chi Minh City were more likely to yield more *Cryptococcus* isolates than sites from Nhon Trach - 92/500 (18%) sites were positive in the city compared with just 15/500 (3%) of sites in Nhon Trach, $P< 0.0001$.

Following the identification of a cluster of 25 clinical isolates (Chapter 2), I performed targeted sampling at addresses of corresponding 25 patients in the cluster. I did not isolate *C. neoformans*, but interestingly 13 out of 25 sites (52%) were positive for non-*neoformans* isolates. In other words, the recovery rate of cryptococci by targeted sampling in the cluster of clinical isolates was 5 times higher than overall random sampling (13/125=10.4 % vs 107/5000=2.1 %, P value <0.001). Moreover, the only positive air samples were obtained in the clinical cluster sites: *C. heveanensis* and *C. laurentii* (Table 3).

Table 3-3 lists strains recovered from targeted sampling. Three *Cryptococcus* species were recovered from 12 tree species, gecko guano, air conditioner filter, and wall swabs. Surprisingly, none of the samples from avian droppings were positive.

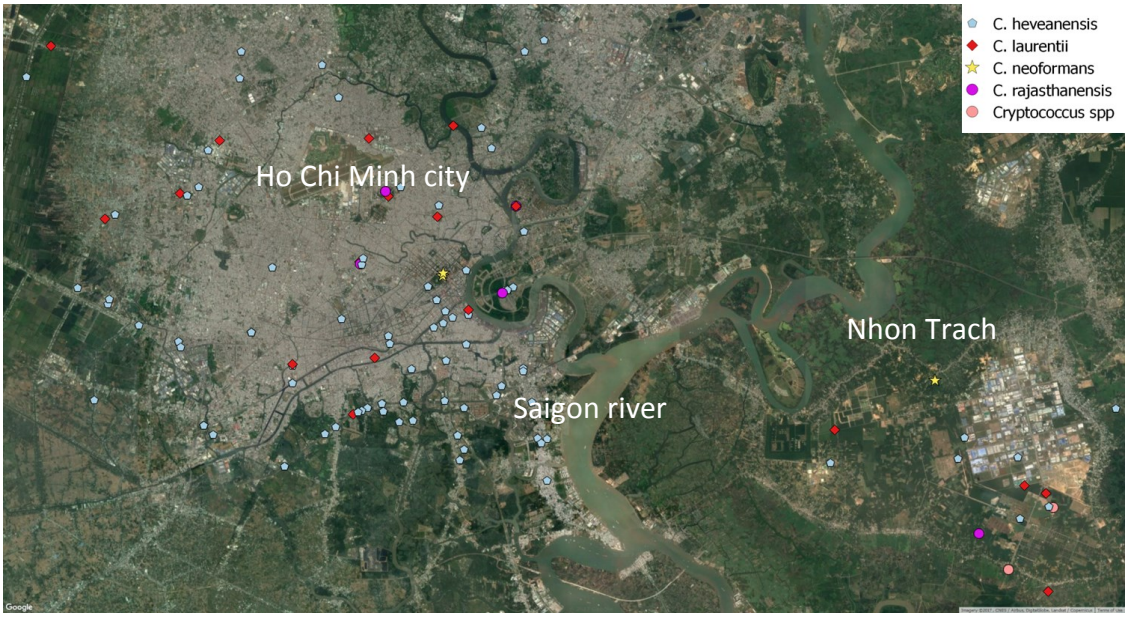


Figure 3-6 Distribution of environmental *Cryptococcus* species in Ho Chi Minh city and Nhon Trach district. These administrative areas are separated by Saigon river. Of 123 isolates, 15 were from Nhon Trach district.

Species	Number of cryptococcal isolates	
	Nhon Trach	Ho Chi Minh city
C. neoformans	1	2
C. laurentii	4	15
C. heveanensis	7	84
C. rajasthanensis	1	4
Undetermined cryptococcus	2	0

Table 3-6 Prevalence of *Cryptococcus* isolates in Nhon Trach and Ho Chi Minh city. All strains from Nhon Trach were isolated by random sampling.

I recovered 5 species of *Cryptococcus* from Nhon Trach (Table 3-5), including the isolation of two unidentified cryptococci and 4 other species. The Nhon Trach sampling area was much smaller than Ho Chi Minh city (143 km^2 vs 484 km^2). In Ho Chi Minh City, I isolated 4 different species. For random sampling, the proportion of *C. laurentii* isolated in Ho Chi Minh city (11 strains, 73%) was significantly higher than that in Nhon Trach (4 strains, 26%), P value = 0.01, 2-sample test for equality of proportions). Similarly, *C. heveanensis* strains were more abundant in Ho Chi Minh city (76 strains, 92%) than in Nhon Trach (7 strains, 8%) (P value <0.001). This disproportion was potentially confounded by sampling seasons and difference in ecology of the two sampling areas. Ho Chi Minh City was mostly sampled in dry season while Nhon Trach was mostly sampled in rainy season.

3.4.4 Clustering of environmental isolates

I identified a cluster of *C. heveanensis* that spans the south eastern districts of Ho Chi Minh city. The cluster included 30 out of 81 *C. heveanensis* strains randomly isolated in Ho Chi Minh City. Districts included in the cluster cover both urban and rural areas and are located in the network of canals and rivers (Figure 3-7). Where the cluster occurred, elevation was generally lower than for the rest of Ho Chi Minh City, and the population density was also lower (Figure 3-8). I had insufficient numbers of *Cryptococcus* of other species to allow a geospatial clustering analysis.

Figure 3-7 Spatial cluster of *C. heveanensis* (radius = 4.84 km, P value = 0.002) was detected in the southeastern Ho Chi Minh city. This cluster spans both urban and rural areas with a network of canals.

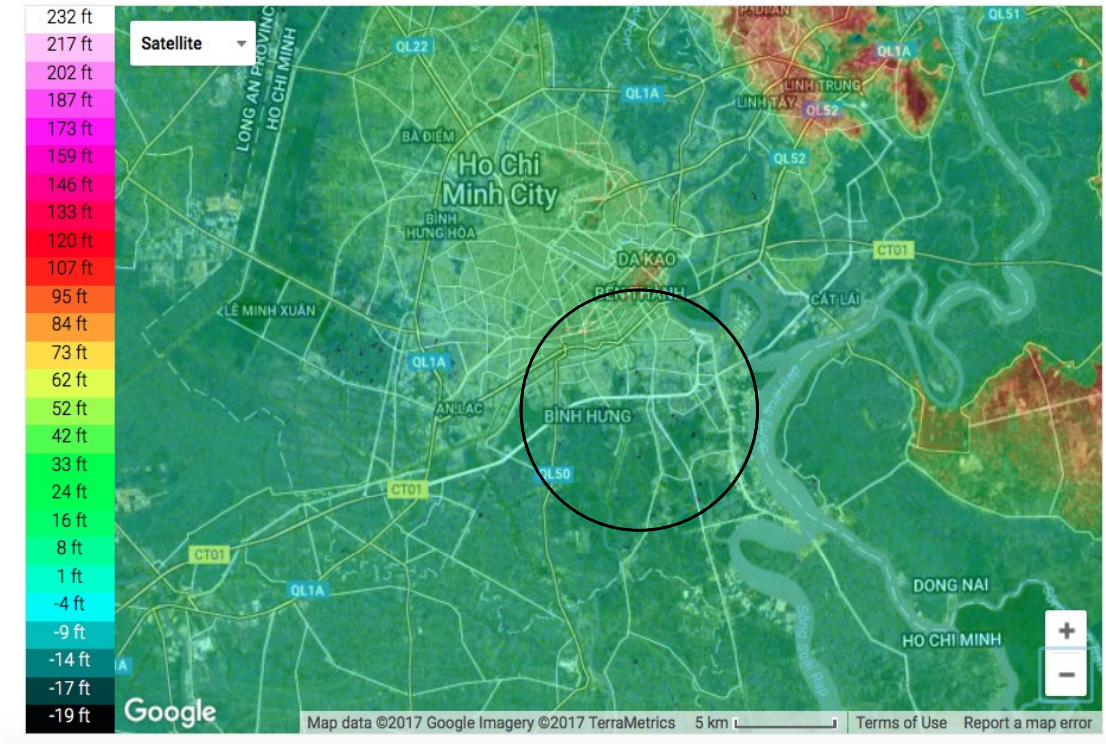


Figure 3-8 Elevation map of Ho Chi Minh city with a spatial cluster of *C. heveanensis*. More than half of the circle (rural areas, dark green) have lower elevation and is less populated than urbanised areas of HCMC (light green).

3.4.5 Association between environmental isolates and arboreal sources

Cryptococcus spp were isolated from 29 of the 79 tree species sampled. *Acacia auriculiformis* was sampled the most frequently, but had the lowest rate of recovery (1.3%, 95% CI 0.16% - 5%), with just two *C. neoformans* strains isolated. For random sampling, the flamboyant tree (*Delonix regia*) harbours the largest number of *Cryptococcus* isolates 19/107 (18%, 95% CI 12% - 26 %) (Table 3-2).

Hopea odorata was associated with 4 different *Cryptococcus* species (*C. neoformans, C. heveanensis, C. laurentii* and *C. rajasthanensis*); however, I lacked power to determine whether it truly held a diverse population. I was unable to demonstrate an association between any particular cryptococcus species and any particular tree species (Figure 3-9).

C. heveanenesis was the dominant environmental population of cryptococci in Vietnam. *C. heveanensis* has also been found to be the most frequently recovered environmental isolate in Cuba, which is somewhat surprising given the geographical distance between each country, and points to a global population dispersal of this species[193]. I isolated *C. heveanensis* from both arboreal and non-arboreal sources. The flamboyant tree (*Delonix regia*) was used as reference to study association between this tree and *C. heveanensis* because flamboyants accounted for the largest frequency of *C. heveanensis* isolates by random sampling (22%, 18/83) in my study. Moreover, this tree was previously documented as an ecological habitat of *Cryptococcus spp* in Cuba[193]. However, my data did not support this association between *C. heveanensis* and *Delonix regia*. They did not reject the null hypothesis that there was no difference in the proportion of flamboyants positive for *C. heveanensis* compared with other trees, with the calculated P value not quite reaching conventional levels of statistical significance (P value = 0.065, Fisher's test).

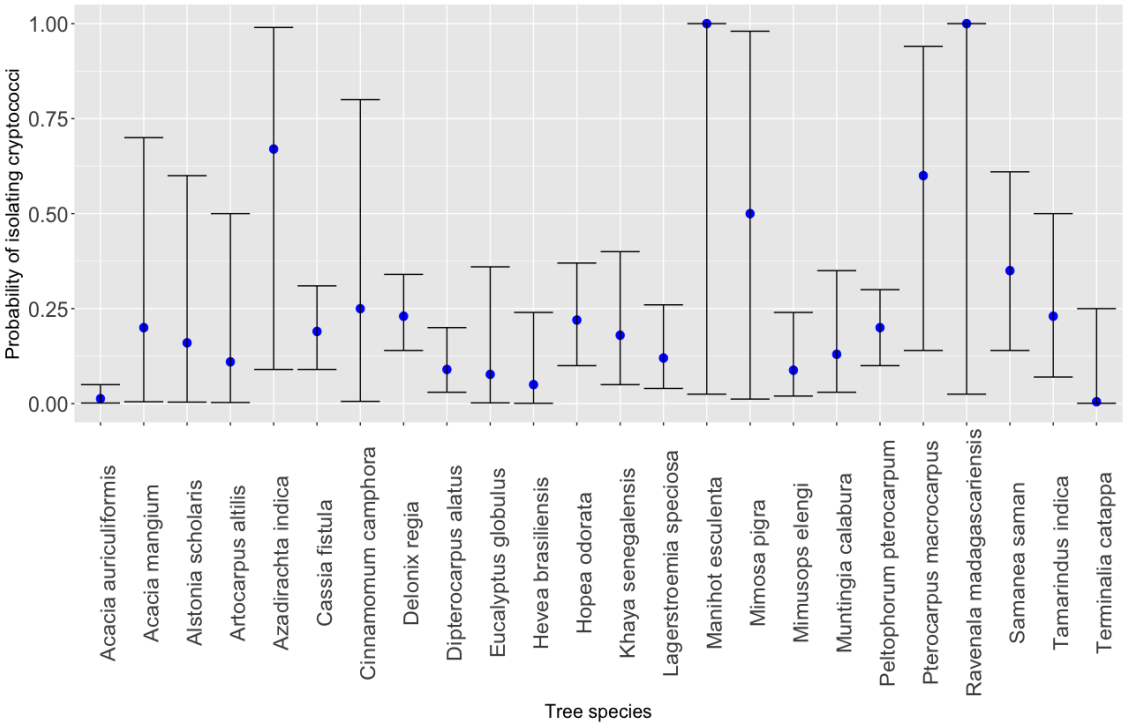


Figure 3-9 Probability of isolating *Cryptococcus spp* associated with tree species from random sampling. 95% confidence intervals for binomial proportion were estimated with Clopper-Pearson method.

3.4.6 Phylogeny of environmental isolates

A maximum likelihood tree was constructed to clarify the phylogenetic relationship of these species based on ITS sequences (Figure 3-10). *C. neoformans* strains clustered with the reference clinical isolate H99. This clustering is supported by 100% bootstrap value, implying environmental *C. neoformans* are potential pathogenic. The main polytomy clade of *C. heveanensis* strains is more closely related to *C. neoformans* than it does with the rest of *Cryptococcus* species. Undetermined *Cryptococcus strains* are more closely related to *C. rajasthanensis* than rare pathogens *C. laurentii*.

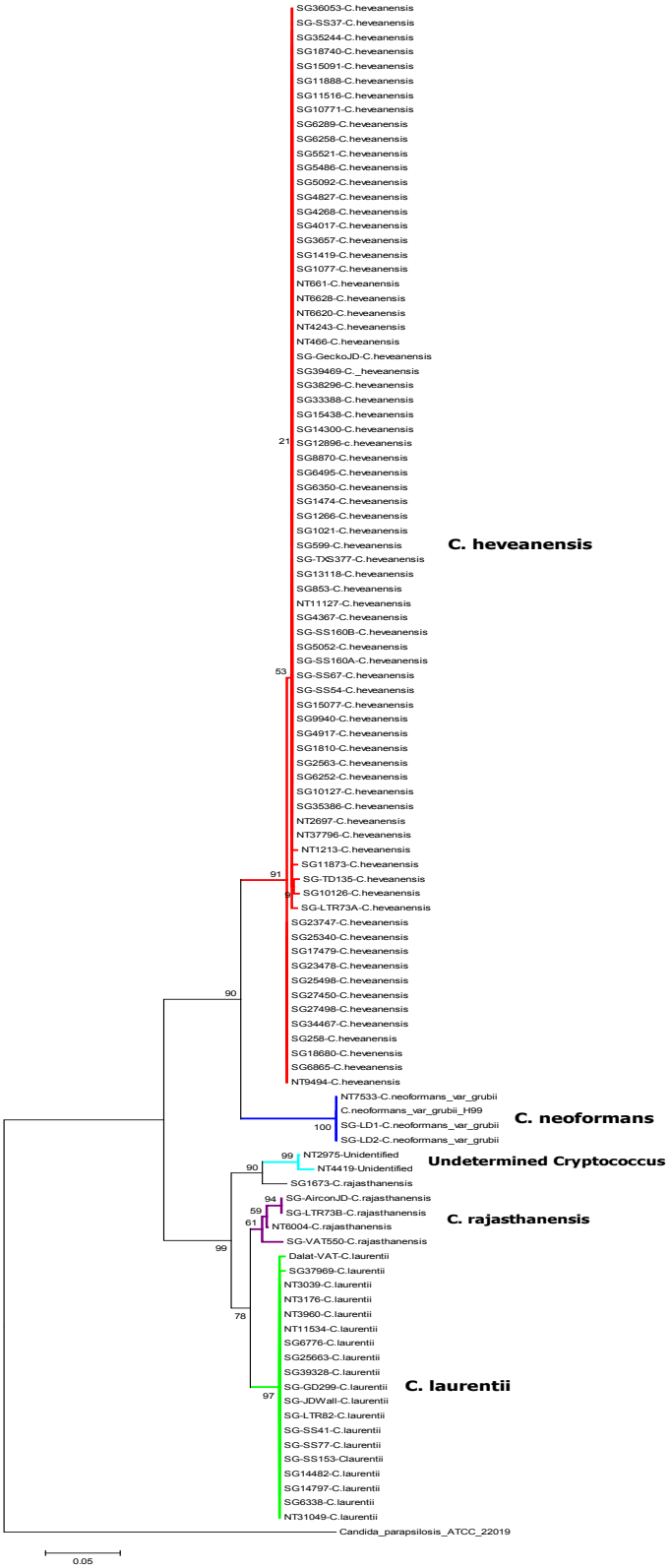


Figure 3-10 Maximum likelihood tree of environmental cryptococcal isolates based on ITS sequence (550 bp) constructed by MEGA 6. *Candida parapsilosis and H99* are outgroup and clinical control, respectively. Numbers on the left of each node represent bootstrap percentages produced by 1000 replications.

3.4.7 Multilocus sequence typing

Multilocus sequence typing confirmed the environmental *C. neoformans* isolates to be sequence type 5 and sequence type 4 (Table 3-7). These genotypes (5 and 4) account for the majority of disease in our clinical practice (97/136, 71 %)[88]. Sequence type 5 and 4 correspond to the AFLP clusters VNIγ and VNIδ, respectively[86].

Strain	CAP59	LAC1	PLB1	GPD1	SOD1	URA5	IGS1	Sequence type
LD1	1	4	2	1	1	5	1	4
LD2	**1**	**5**	**2**	**3**	**1**	**1**	**1**	**5**
NT7533	1	5	2	3	1	1	1	5

Table 3-7 MLST profiles of three environmental *C. neoformans* var*. grubii* isolates

3.4.8 Characterisation of virulence-associated phenotypes

3.4.8.1 Urease production

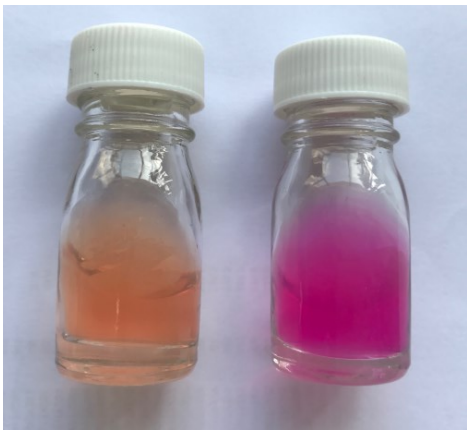


Figure 3-11 Urease Christensen agar test to detect cryptococci. Negative control (orange tube): *Candida parapsilosis*. Positive: *Cryptococcus spp*.

Urea Christensen agar was used for tentative identification of cryptococci [135]. All *C. neoformans* isolates turned pink, usually within 24 hours of inoculation. Other *Cryptococcus* species (120 isolates) gave positive tests 48 hours post-incubation at 30 ^{o}C, suggesting their urease activity is lower than that of *C. neoformans* (Figure 3-11).

3.4.8.2 Melanin formation

I used L-DOPA agar to characterize the melanin production of the environmental *Cryptococcus* species. All 3 *C. neoformans* isolates produced melanin resulting in the

characteristic dark brown colonies, whereas only 7/19 (37 %) of *C. laurentii* strains exhibited melanisation by 7 days. Moreover, the colonies of non-*neoformans* melanin-producing isolates were lighter than those of *C. neoformans*, indicating less vigorous melanin production (Figure 3-12). Only one *C. heveanensis* isolate exhibited melanisation attribute (Table 3-8 & 3-9). While *C. rajasthanensis* melanized on L-DOPA media, it was unable to grow at 37^{o}C, implying it is unlikely to be able to cause disease in humans.

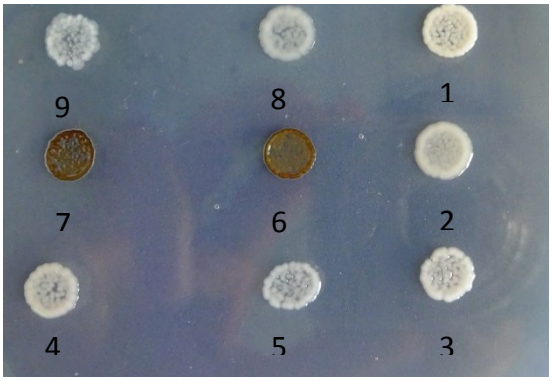


Figure 3-12 Melanin formation of *Cryptococcus spp* in L–DOPA agar. 1&2: *C. laurentii*, 3: *C. rajasthanensis, 4&5: C. heveanensis. 6&7: C. neoformans*. 8&9: undetermined *Cryptococcus*.

Isolate	% isolates in each species exhibiting particular phenotypes		
	30^{o}C	**37^{o}C**	**Melanin**
C. neoformans* var. *grubii	100	100	100
C. laurentii	100	47	42
C. heveanensis	100	1	1
C. rajasthanensis	100	0	17
Unidentified *Cryptococcus*	100	0	0

Table 3-8 Percentage of isolates can grow above 30^{o}C in YPD or produce melanin on L-DOPA agar

ID-Species	Temperature		Melanin expression
	37 °C	39 °C	
LD1-*C.neoformans* var. *grubii*	+++	+	++
LD2-*C.neoformans* var. *grubii*	++	-	++
NT7533-*C.neoformans* var. *grubii*	+++	++	++
NT3176-*C.laurentii*	+++	-	+
SG14482-*C.laurentii*	++	-	-
SG25663-*C.laurentii*	++	-	-
SG31049-*C.laurentii*	+	-	-
SG39328-*C.laurentii*	++	-	-
SG6776-*C.laurentii*	++	-	+
SGLTR82-*C.laurentii*	+	-	-
SGWallJD-*C.laurentii*	++	-	-
DLVAT-*C.laurentii*	-	-	+
NT3039-*C.laurentii*	-	-	+
SG299-*C.laurentii*	-	-	+
SGSS153-*C.laurentii*	-	-	+
SGSS41-*C.laurentii*	-	-	+
SG33388-*C.heveanensis*	+	-	-
SG4827-*C.heveanensis*	-	-	+
SG-LTR73B-*C.rajasthanensis*	-	-	+

Table 3-9 Strains exhibit ability to grow above 37°C in YPD agar or to produce melanin in L-DOPA agar. +++: strongly grow, ++: grow well or produce melanin robustly, +: fairly grow or fairly produce melanin, - : fail to survive or fail to produce melanin

3.4.8.3 Thermotolerance

Susceptibility to mammalian temperature is key to pathogenicity. I exposed all isolates (*C. neoformans*, *C. laurentii, C. heveanensis, C. rajasthanensis* and undetermined *Cryptococcus*) to a range of temperature (30°C, 37°C, 39°C). All strains

had been initially selected through growth at 30°C (a typical shaded temperature in the south of Vietnam), and thus all grew well at this temperature following purification. However, at 37°C the *C. neoformans* isolates grew more robustly than the rest. Only one isolate of *C. heveanensis* could grow at 37°C. Potentially pathogenic isolates (*C. neoformans & C.laurentii*) can withstand 37°C and above better than other species. However, only *C. neoformans* (3 isolates) could survive at 39°C (Figure 3-13). Interestingly, of 8 *C. laurentii* growing at 37°C, 7 strains were derived from Ho Chi Minh city and only one from Nhon Trach. The ability of *C. laurentii* strains to survive at elevated temperature was not associated with melanin production (P=0.63, Fisher's exact test).

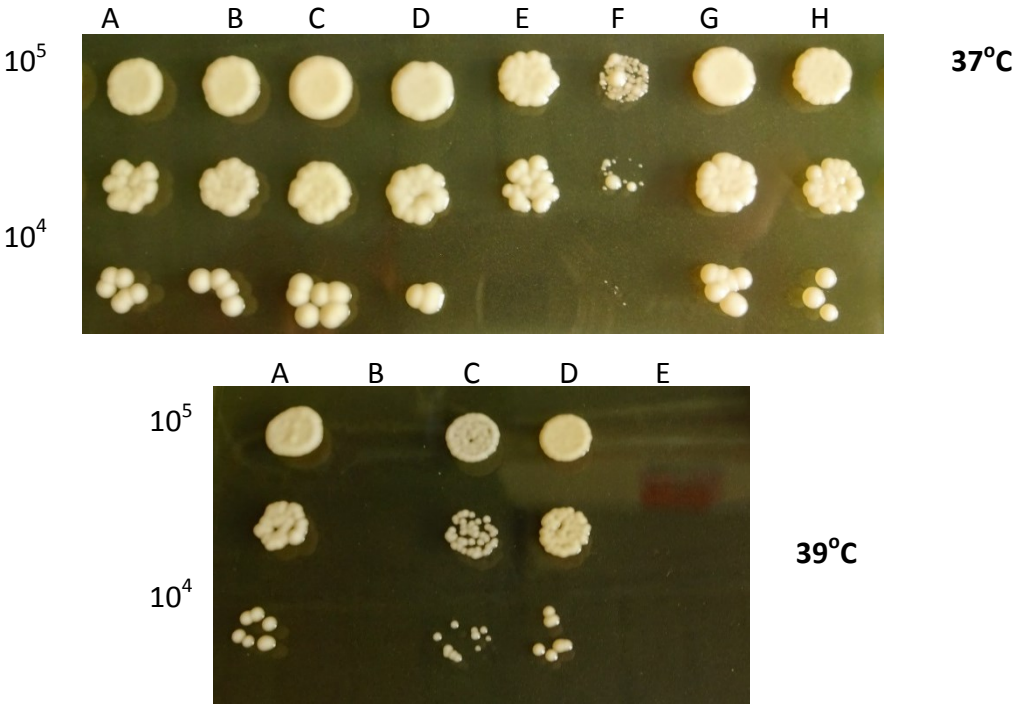


Figure 3-13 Temperature-dependant growth of cryptococci at 37°C (upper panel) and 39°C (lower panel). At 37°C, *C. heveanensis* (E) and *C. laurentii* (F); the rest are C. neoformans (A,B,C) and *C. laurentii* (D,G,H) that grew equally well. At 39°C, C. *neoformans* (A, C, D); *C. heveanensis (B)* and *C. laurentii (E)*

3.4.8.4 Capsule size

The capsule size of the *C. neoformans* isolates was significant larger than that of all other *Cryptococcus* species post-incubation in capsule-inducing agar (DMEM) (P value < 0.01, multiple comparison with P value adjustment performed by Hochberg

method). *C. neoformans* also exhibited the most variation in capsule size compared with other species (Figure 3-14).

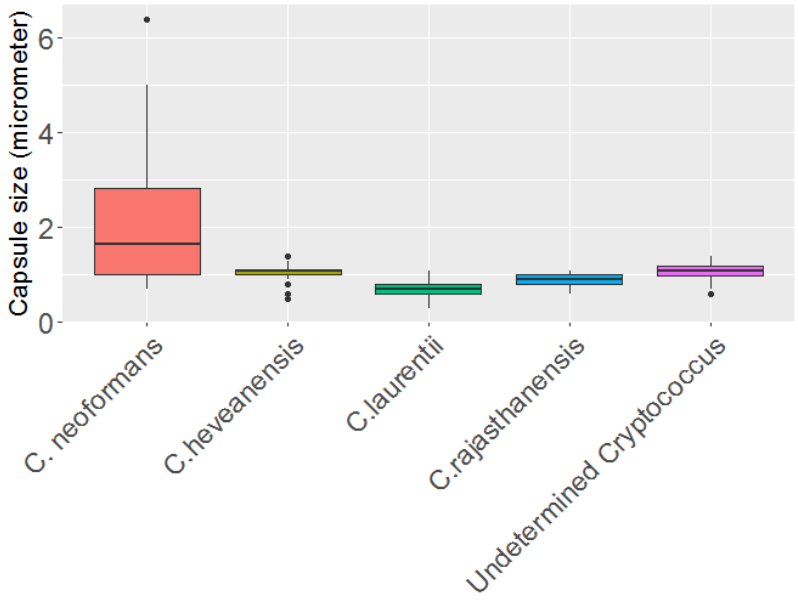


Figure 3-14 Variation in capsule size between environmental *Cryptococcus spp*. 3 strains of each species were measured for capsule size. 40 cells were measured per species

3.4.8.5 *Ex vivo* CSF survival

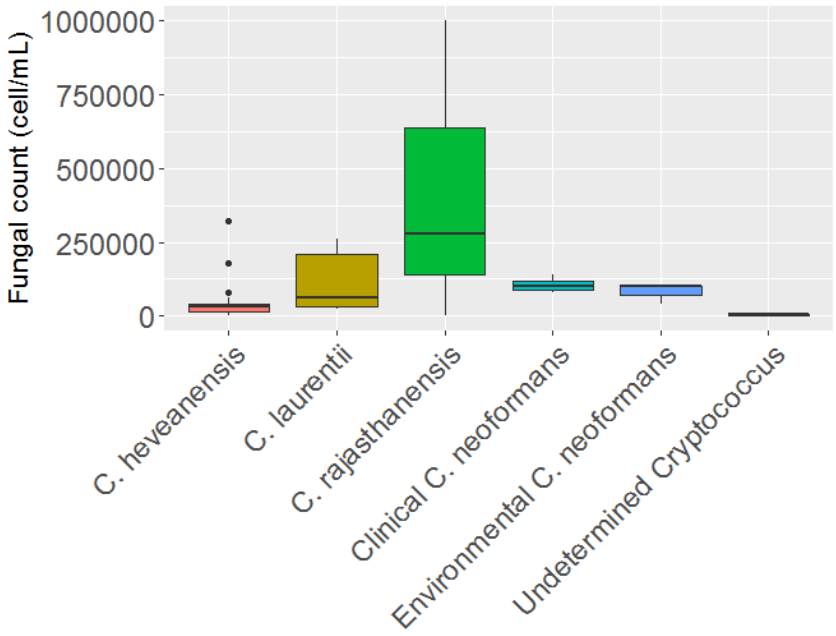


Figure 3-15 Fungal count of *Cryptococcus* isolates after 2 days of culture in *ex vivo* CSF. *C. heveanensis* (n=33), *C. laurentii* (n=8), *C. rajasthanensis* (n=3), environmental *C. neoformans* (n=3), clinical *C. neoformans* (n=3), Undetermined *Cryptococcus* (n=2).

The ability to grow within the cerebrospinal fluid (CSF) is a pre-requisite for infection within the central nervous system, and specific genes associated with survival in the human host CSF have been identified [201]. The ability to grow in CSF is considered a virulent characteristic. In order to define the virulence potential of the environmental isolates, I grew them in pooled human CSF. I used *C. neoformans* isolated from patients within our hospital as controls. I found significant differences in fungal counts between environmental *C. neoformans* vs *C. heveanensis* (P=0.03, Wilcoxon rank sum test), and clinical *C. neoformans* vs *C. heveanensis* (P=0.015). Surprisingly, *C. rajasthanensis* strains had both the most rapid growth resulting in the highest fungal burdens (Figure 3-15), and the highest variability in fungal counts, suggesting this species has a somewhat heterogenous population.

3.4.9 Survival analysis of *Cryptococcus spp* using *Galleria mellonella* infection model

To compare the survival of *Cryptococcus spp*, I randomly selected two strains of each species: *C. neoformans, C. heveanensis, C.laurentii, C. rajasthanensis* and unidentified *Cryptococcus* for larva killing assays at 30°C. The virulence of environmental strains of *C. neoformans* was not significantly different from the other species (Figure 3-16). *C. heveanensis* appeared more virulent than *C. laurentii* (P value=0.01)

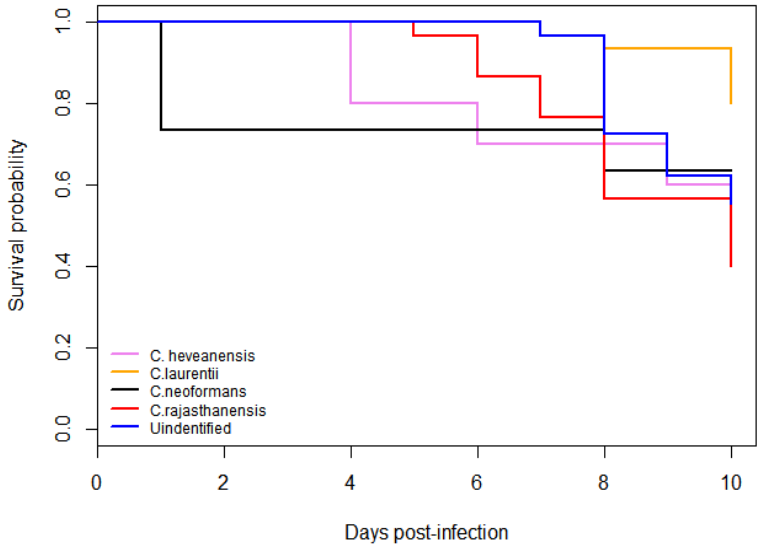


Figure 3-16 Survival curves of five environmental *Cryptococcus* species infected with *Galleria mellonella* at 30°C. P value >0.05 obtained from the Cox proportional-hazards model for comparison between *C. neoformans* and others. Each species was represented by two isolates (15 larvae/isolate). N=30 larvae/arm

Human disease requires growth at 37°C. Thus I repeated the larva infection assays at this temperature. I used only *C. neoformans* and *C. laurentii* isolates because they can grow reliably at 37°C. Each species was represented by two isolates. Larvae challenged with *C. laurentii* survived significantly longer than those infected with either clinical or environmental isolates of *C. neoformans* (Figure 3-17). In fact, most of the larvae survived beyond the ten-day observation period, consistent with *C. laurentii* being less virulent than *C. neoformans* at 37°C.

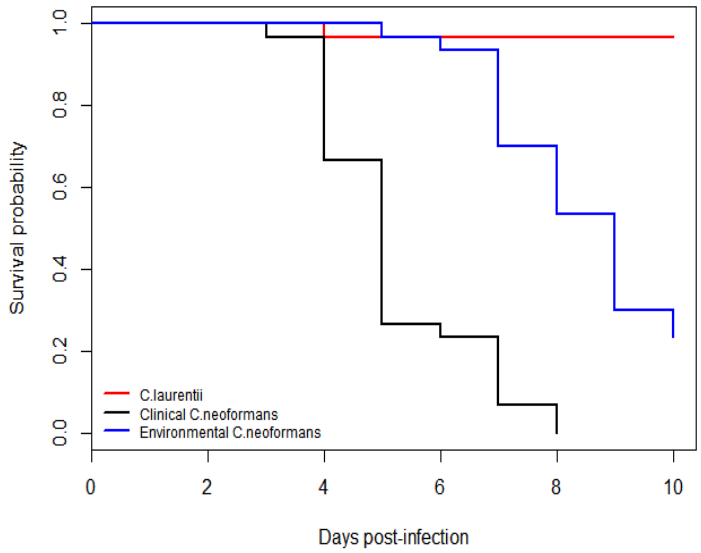


Figure 3-17 Survival analysis of *C. neoformans* vs *C. laurentii* infected with *Galleria mellonella* at 37°C. P value < 0.001 obtained from the Cox proportional-hazards model for comparison between *C. laurentii* and others. Each species was represented by two isolates (15 larvae/isolate). N=30 larvae/arm

3.4.10 Characterisation of underdetermined *Cryptococcus* species

During random sampling in Nhon Trach, I isolated two isolates of underdetermined *Cryptococcus* species from *Manihot esculenta* and *Mimosa pigra* (Table 2 and Figure 3-6). ITS-based phylogeny reveals they are closely related to *C. rajasthanensis* (Figure 9). Unlike *C. neoformans*, they did not exhibit urease activity on Christensen's agar or melanization when grown on L-DOPA agar (Figure 3-11 & 3-12). Furthermore, like the majority of non-medically significant *Cryptococcus* , they were unable to grow at 37°C. Their production of capsule and ability to survive *ex vivo* CSF were statistically significantly different to *C. neoformans* (Figures 3-14 and 3-15). Taken together, these suggest that these *Cryptococcus* of undetermined species are extremely unlikely to be able to cause disease in humans.

3.5 Discussion

3.5.1 Environmental niches of *Cryptococcus spp*

I hypothesized that *Cryptococcus spp* occupy particular ecological niches. Overall I isolated 123 environmental cryptococci isolates, with 107 isolates obtained from random sampling and 16 derived from targeted sampling (Table 3-2 & 3-3). I isolated *Cryptotoccus* species from a wide variety of arboreal and non-arboreal sources. Although association between tree species and *Cryptococcus* was not established, *Hopea odorata* (Cengal Pasir, Vietnamese: sao đen), a native plant in southern Vietnam, is an interesting host for *Cryptococcus* species, harbouring *C. neoformans* and three other non-*neoformans* cryptococci species (Table 3-2 & 3-3). *Hopea odorata* is native to Southern Asia and Indochina countries, including India, Bangladesh, Myanmar, Thailand, Laos, Vietnam and Malaysia. This is the first time isolation of *C. neoformans* from this tree has been reported.

I found the prevalence of *Cryptococcus* species in the urban area (Ho Chi Minh city) was significantly higher than in rural areas (Nhon Trach district). However, it is possible that this result is confounded by the season of sampling. I sampled the rural areas mostly in the rainy season (from April to August) and the urban areas in the dry season (from November to January). The recovery of environmental isolates of *C. neoformans* has been reported to be more efficient in the dry season in some tropical countries including Thailand, Zambia and India [173] [57] [192] . However, the contrary observation has also been reported in similar environments - in Colombia *C. neoformans* was recovered more frequently in the rainy season [172] [204] . I was surprised that I failed to recover *C. neoformans* from avian sources (bird, pigeon, chicken, duck guano and soil contaminated with avian droppings) although these sources were common habitats of *C. neoformans* as previously reported in Africa, Thailand, India, USA, Columbia [26,56,57,171,205]. More *Cryptoccus species* are found in Nhon Trach although the sampling area is much smaller than Ho Chi Minh city (143 km^2 vs 484 km^2). 12 out of 15 isolates in Nhon Trach were derived from samples collected in the field where trees predominate (rural areas) (Figure 3-6).

My data showed that non-*neoformans* strains were ubiquitous and were able to be isolated from a wide spectrum of sources including trees, guano, soil, air, dust (air

filter, inanimate objects such as a wall). The recovery rate of *C. neoformans* in my study was low (0.06%, 3/5125). The literature suggests that the percentage of positive environmental samples can vary from 0 to 10 % with a likely higher rate of isolation in endemic areas [56] [52] [7]. I used randomized sampling, which can explain my relatively low recovery rate of *C. neoformans* (0.06%, 3/5000). I may simply have missed the most important ecological niche for *C. neoformans* in Vietnam, although at each sampling point I used a diverse approach to harvest samples including air, soil, and tree swabs. One practical problem was the size and inaccessibility of branches of trees in Vietnam – I mostly had to sample the lower trunks of trees, so could have missed *Cryptococcus* that was colonising upper trunks/branches or even in flowers. In Thailand, 2 % of environmental *C. neoformans* was isolated in eucalyptus flowers[59]. A further possibility is that environmental *C. neoformans* in Vietnam produce melanin poorly, making it difficult to detect them on birdseed agar. While melanin production is usually obvious in Vietnamese clinical isolates, Pereira *et al* compared the melanin intensity of 15 clinical vs 15 environmental *C. neoformans* and *C. gattii* and found a wide variation in melanin production in environmental isolates, which may reflect their different exposure to environmental stresses[206]. A further explanation for my low recovery of C. *neoformans* is that it simply grows very slowly in the environment, and is mainly quiescent, and my culturing techniques were insufficiently sensitive to detect it. A final explanation could be failure to detect because of overgrowth of filamentous molds post-inoculation. However, I processed samples as soon as possible, mostly within 2 days of collection to minimize this risk. Moreover, mold overgrowth did not seem to impair the growth of other *Cryptococcus* species. The methods I used were consistent with Litvintseva *et al's* validated protocol; like me, they processed samples within 3 days of collection[171].

3.5.2 Speciation of *Cryptococcus* species

According to the manufacturer Biomerieux, the ID32 C system covers 62 yeast taxa including 7 *Cryptococcus* species which are medically important. A previous study using this system identified correctly 87% of *Candida* strains [207]. Here, I found the ID32 C only correctly identified only 14% of *Cryptococcus spp* (Table 3-5). However,

this method can detect 100% of *C. neoformans* in agreement with ITS sequencing (Table 3-5). The misidentification of *C. heveanensis*, *C. rajasthanensis* and the unidentified *Cryptococcus species* is likely a consequence of these species not being included in the ID32 database, and a result of a 'best fit' approach. Efficiency of ID32 C was also lower than Vitek MS[208] for identification of a variety of clinical yeasts. My finding suggests ID 32C can be used for screening presumptive *C. neoformans*, but additional confirmation should be followed by culture on bird seed agar, a selective medium for *C. neoformans,* and ITS sequencing.

In terms of accuracy, the sequence-based method, e.g. ITS barcode, in combination with phenotypic assays was superior to ID32 C alone. The former should be employed along with phenotypic assays because of the (small) possibility of DNA contamination from other fungal species. This has been reported to occur (during DNA extraction or PCR assay) as frequently as 3.3 % in negative controls and 4.7 % in PCR mixture[209]. Commercial products such as zymolyase powder, PCR buffer, TaqMan probes and master mix solutions can contain traces of fungal DNA[210].

In contrast to *C. neoformans*, *C. laurentii* is a rare causative agent of opportunistic infections in humans. *C. laurentii* mostly infects skin, bloodstream and the central nervous system. Patients with impaired cell-mediated immunity, such as HIV infection (CD4 count<100), organ transplant, and immunosuppressant use, or patients with exposure to azoles or invasive catheter devices are predisposed to *C. laurentii* infection [47 211].

C. heveanensis was the most abundant species (76%) in my environmental cryptococcal population. *C. heveanensis* has a diverse global distribution, having previously been recovered from trees and cacti in Havana, Cuba [193], tropical forests in Thailand, and glacial meltwater in Argentina[212 213]. However, human disease due to this species has not been reported because they are not frequently tolerant to human temperature.

Another non-pathogenic *cryptococcus* is *C. rajasthanensis*, which was first isolated from tree flowers in Rajasthan state, India[214] and later from hollow trees in Brazil, Thailand, and China[215]. This isolate is closely related to *C. laurentii* using ITS sequence analysis as previously reported[215].

I also isolated an unidentified *Cryptococcus* species, whose ITS sequence was deposited in the Genbank but its phenotype was not characterised previously.

3.5.3 Spectrum of tree species hosting *Cryptococcus spp*

I sampled 79 tree species - somewhat more than the number examined in other studies (for example, India (17 species) or Colombia (35 species))[216] [172]. Nearly 40% (29/79) of tree species that I examined were positive for *Cryptococcus spp*. A number of tree species that I sampled in Vietnam have been shown to harbour *C. neoformans* and *C. gattii* in India, including *Cassia fistula*, *Mimusops elengi*, *Azadirachta indica*, *Tamarindus indica* and *Alstonia scholaris*[26] . However, I found only non-neoformans species on these trees in my study (Table 3-2). Notably, the tropical almond tree *Terminalia catappa* and rain tree *Samanea saman* host non-*neoformans* cryptococci in Vietnam but have been shown to host the *C. neoformans* species complex in Colombia[172]. *Terminalia catappa* was notable as this tree is associated with *C. gattii* serotype C in Colombia[172]. I found no *C. gattii* in my sampling although we see 5-10 cases of this infection each year at the Hospital for Tropical Diseases, HCMC.

My study is the first time that *C. neoformans* has been isolated from *Acacia auriculiformis* and *Hopea odorata*. A*. auriculiformes* has previously been associated with fungi of importance to humans – it is the ecological niche of rust fungi of the genus *Endoraecium spp* in Australia[217]. A number of my *Cryptococcus* isolates were associated with *Cassia fistula*. This species is an important host of *C. neoformans* in India [216] although I isolated only non-*neoformans* species from this tree. Similarly, while 15 % of *Mimusops elengi* samples were found to be colonized with *C. neoformans* in India, I had a recovery rate of 12% of non-*neoformans Cryptococcus* from this tree in Vietnam (Tables 3-2 & 3-3). It is not clear why I had such low rates of recovery of *C. neoformans* from these trees given the Indian experience. Possible explanations include different sampling sites from the trees in the different studies, different sensitivities within the laboratories, or failure of reporting of non-medically important species from the published studies. However, it is likely that the botanical niches of *C. neoformans* vary from country to country.

Figure 3-18 The 30 April park in Ho Chi Minh city, with predominance of golden oak (*Hopea odorata*), where two environmental *C. neoformans* strains were isolated.

Hopea odorata is of great interest, with a wider variety of species of *Cryptococcus* in my study. 24 % of samples from this tree are positive with both *C. neoformans* and three other *Cryptococcus* species (Table 3-2 & 3-3). Were I to have the resources to repeat my study, I would focus on this species. However, it represents particular challenges to sample, being exceptionally tall with height commonly greater than 45m and a paucity of low branches (Figure 3-18).

3.5.4 Clustering of *Cryptococcus spp*

I found a spatial cluster of randomly isolated *C. heveanensis* strains in Ho Chi Minh city. This cluster overlapped with the cluster of cases of meningitis due to ST5 strains described in Chapter 2. Interestingly, the recovery rate of non-*neoformans* cryptococci in the cluster sites of clinical *C. neoformans* ST5 was five times higher than that of random sites, and two air samples of cluster sites contained *non-neoformans* cryptococci. Taken together, these data suggests the ecological niche of *C. neoformans* and non-*neoformans* isolates overlaps in Vietnam.

Given that cryptococcal meningitis occurs with reasonable frequency in Vietnam, I was surprised (and disappointed) not to recover more isolates of *C. neoformans*. Discrepancies between the predominant clinical cause and the apparent dominant environment isolate have been documented in other studies. Extensive sampling in Zambia revealed that the majority of environmental isolates were of the VNB group (83%), whereas VNI strains predominate in human infection (76%) [173]. A study from Bostwana similarly found that VNI strains accounted for 30% of clinical isolates but only 6.5% of the environmental population[218]. Ecological sampling in Havana, Cuba recovered only one *C. neoformans* out of 660 samples [193]. A recent study from South Africa performed in the densely populated townships around Cape Town where human disease is common recovered only 2 *C. neoformans* isolates from 248 samples collected from trees, soil and pigeon guano [219]. A possible explanation for the low recovery rate of *C. neoformans* from the environment may be that the human adapted species have low rates of growth in the environment – other species are better adapted to growth as saprophytes on trees.

3.5.5 Phenotypic assays of *Cryptococcus spp*

I tested environmental isolates of different species using standard virulence assays including temperature-dependent growth, melanin production, capsule size, and urease activity to determine whether such environmental isolates could potentially cause human disease. As expected, *C. neoformans* isolates had the most robust expression of such phenotypes. My data suggest that thermotolerance is probably the single most important virulence factor. All species grew similarly at 30^{o}C; at 37^{o}C there was a fall in the viability of *C. heveanensis* and some *C. laurentii* strains; at 39^{o}C, the growth of both *C. heveanensis* and *C. laurentii* was markedly inhibited. *C. neoformans* strains, both clinical and environmental in source, grew well at all temperatures.

This variability in thermotolerance was also manifest in the *Galleria* model. At 30^{o}C, the virulence of *Cryptococcus spp* was indistinguishable. However, incubation at 37^{o}C revealed differences in virulence between *Cryptococcus neoformans* and *Cryptococcus laurentii*. Clearly the ability to grow at 37^{o}C is critical for human infection and a good indicator of the likelihood that a species can be a human

pathogen. A recent case report described cryptococcal meningitis in an HIV patient in Guatemala, due to infection with *Cryptococcus liquefaciens* complicated by co-infection with *Mycobacterium tuberculosis*[220]. *C. liquefaciens* was previously considered as non-pathogenic yeast because of its poor growth at mammalian temperatures. Moreover, this species did not produce melanin in L-DOPA agar. Its capsule was visible using India ink staining and survival of *G. mellonella* infected with *C. liquefaciens* was comparable to larvae infected with *C. neoformans* at 25^{o}C [221]. It is feasible that rises in average global temperatures could drive the development/selection of thermotolerance in environmental fungal species which may have a bystander effect of increaseing their pathogenic potential[222]. In my environmental *cryptococcus* population, it was interesting to see that there was evidence of variability within species - 1 % of *C. heveanensis* and 47 % of *C. laurentii* strains could grow at 37^{o}C. This is consistent with the occasional reports of *C. laurentii* as a human pathogen[211]; I am unaware of any reports of disease due to *C. heveanensis*, but presumably this is a possible rare event.

In this study, I found that *C. neoformans* was more virulent than *C. laurentii* in the *Galleria* larvae although both species grew well at 37^{o}C, given that *C. neoformans* strains were significantly more melanised than *C. laurentii* when grown on bird seed agar and L-DOPA agar. So far there has been no study investigating correlation between virulence of *C. neoformans* in *G. mellonella* and pathogenicity in humans. This suggests the important role of melanisation in cryptococcal virulence. *G. mellonella* infection model was previously used to identify the hyperpathogenic *C. gattii* strains causing Vancouver Island outbreak (VGIIa) and other major molecular types (VGI, VGII, VGIII and VGIV)[166]. A number of *C. gattii* strains (VGI, VGII, VGIII and VGIV) were significantly less virulent than the highly pathogenic Vancouver Island outbreak strain CDCR265 (VGIIa) [166].

Finally, the failure of my novel species to grow at 37 or 39^{o}C suggests that it is unlikely to be able to cause human disease.

3.6 Conclusion

In conclusion, I isolated 123 cryptococci in the environment with 22 potential pathogens (3 *C. neoformans isolates* and 19 *C. laurentii isolates*). I did not identify

species-specific niches because I lacked power for *C. neoformans*. However, the association between increased recovery rates of non-*neoformans Cryptococcus* species from sites of clinical disease suggests that habitats for different species are highly similar and would require high resolution ecological techniques to tease apart. Of note, I could not associate any particular tree species with cryptococcal species. More sensitive environmental recovery techniques are needed to better understand the ecology of *Cryptococcus* species of medical importance.

Chapter 4

Understanding pathogenicity of *C. neoformans* var. *grubii* using the *Galleria mellonella* infection model

4.1 Introduction

We previously reported that cryptococcal meningitis occurring in HIV-uninfected patients in Vietnam was caused almost exclusively by isolates from an AFLP-defined cluster of *C. neoformans* var. *grubii* which we named VNIγ [86]. We subsequently showed that this cluster corresponded precisely with multi-locus sequence type 5 (ST5) [88]. The ST5 strain predominates in China, and neighbouring countries such as Japan, South Korea and Vietnam, where disease in HIV uninfected patients has been frequently reported [169] [223] [224]. Given this host distribution, I hypothesize that ST5 strains have increased pathogenicity – ability to cause disease - compared with non-ST5 strains. To explore this in the laboratory setting requires a suitable model of infection. The mouse model of cryptococcal infection is well established but has some significant disadvantages in that it is expensive, involves use of a sensate vertebrate, and is not practical for higher throughput applications [160].

More than one decade ago, the larvae of the wax worm *Galleria mellonella* were first employed to study the virulence of *Candida albicans* [164]. The demonstration of correlation between the virulence of this clinically important pathogen in *G. mellonella* with that in mice was established, which triggered the widespread use of this model in other fungal pathogens including *Candida krusei, Aspergillus fumigatus, A. flavus, C. neoformans* and *C. gattii* [166] [162]. The popularity of the *G. mellonella* model is driven by multiple advantages over mammalian hosts. The larvae are cheap to grow and maintain, suitable to high-throughput studies and there are no current ethical concerns. The larval immune system is similar to the innate immune system of vertebrates with both humoral and cellular responses. Humoral immune responses of the insect include melanization and production of antimicrobial peptides. The cellular components consist of hemocytes, which function similarly to macrophages and neutrophils in vertebrates [162] [225].

The validity of the *G. mellonella* infection model in interrogating the pathogenesis of human cryptococcal disease is supported by concordance between the *Galleria* and mouse infection models. Mylonakis *et al* was the first to use this model to study cryptococcal virulence, host response to infection and *in vivo* susceptibility to antifungals. They found that several cryptococcal genes involved in mammalian virulence are also important in in larval infection, including *CAP59, GPA1, RAS1* and *PKA1* [98]. They also found the antifungal combination of amphotericin B and flucytosine was significantly more effective than amphotericin B alone in treating infection in *Galleria*. This is consistent with data from humans where the combination has been associated both with better rates of fungal clearance and reduced mortality.[67]

Liu *et al* used the *Galleria* model to perform systematic genetic analysis of virulence genes in *C. neoformans*. Using knockout mutants they discovered novel genes regulating melanization and capsule production and identified a capsule-independent antiphagocytic pathway regulated by Gat201 transcription factors[226]. *G. mellonella* killing assays have also contributed significantly to investigating multiple kinase and transcription factor genes associated with virulence [227] [34].

The morphogenesis of *C. neoformans* has previously been studied using the mouse lung infection model[228]. Zaragoza *et al* demonstrated that the *Galleria* model can also be used for this purpose, with capsule enlargement occurring within 2 hours post-infection, and induction of a sub-population of giant (Titan) cells accounting for up to 5 % of all cells by 4 days post-infection[167].

In addition to significant work interrogating virulence using type strains, the *Galleria* model has been used to compare the virulence of the major molecular types of *C. gattii*, notably the highly pathogenic Vancouver Island outbreak strain VGIIa[166]. Firacative *et al* found variation in virulence between *G. gattii* major molecular types (VGI, VGII, VGIII and VGIV); perhaps surprisingly the Vancouver outbreak strain VGIIa was not the most virulent. Together, these studies support the versatility and validity of the *G. mellonella* infection model in elucidating cryptococcal pathogenesis. However, the relative virulence of different clinical isolates of *C. neoformans* var. *grubii* of different lineages is largely unstudied in model systems –

most studies of virulence focus on the H99 type strain. The purpose of this part of the thesis was to use the *Galleria* infection model to interrogate the relative virulence of different lineages of *C. neoformans* from Vietnam, because of the different host distributions we have observed in clinical practice.[86 88]

4.2 Aims

The aims of the chapter are to use the *Galleria mellonella* infection model to determine whether it can detect differences in the virulence of isolates of *C. neoformans* according to their lineages and source (HIV uninfected or infected patients, clinical or environmental), and between Vietnamese *C. gattii* and *C. neoformans* ST5 strains.

4.3 Methods

4.3.1 Strains

Clinical isolates: The isolate characteristics are presented in Table 1. Strains were randomly selected from those archived within the mycology research laboratory at OUCRU. Isolates from HIV infected patients were from patients enrolled into a randomized controlled trial of antifungal therapy performed at the Hospital for Tropical Diseases [67]. Strains from HIV uninfected patients were randomly selected from those from a prospective descriptive study of CNS infections performed at the Hospital for Tropical Diseases, Ho Chi Minh City [180].

Environmental isolates: Environmental isolates were the 3 *C. neoformans* isolates described in chapter 2 - LD1, LD2 and NT7533.

Typing

All isolates were typed using MLST described in the chapter 2. The STs are described in Table 1.

C. gattii strains: BMD873, BMD1377, BMD800 isolated from HIV-negative patients enrolled in BMD study[180]

ID	Source	Sequence type
BK02	HIV patient	ST4
BK111	HIV patient	ST4
BK120	HIV patient	ST4
BK14	HIV patient	ST4
BK151	HIV patient	ST4
BK156	HIV patient	ST4
BK182	HIV patient	ST4
BK193	HIV patient	ST4
BK224	HIV patient	ST4
BK225	HIV patient	ST4
BK30	HIV patient	ST4
BK35	HIV patient	ST4
BK48	HIV patient	ST4
BK56	HIV patient	ST4
BK57	HIV patient	ST4
BK59	HIV patient	ST4
BK73	HIV patient	ST4
BK74	HIV patient	ST4
BK89	HIV patient	ST4
BK90	HIV patient	ST4
BK42	HIV patient	ST5
BK139	HIV patient	ST5
BMD101	HIV-uninfected patient	ST5
BMD1198	HIV-uninfected patient	ST5
BMD1228	HIV-uninfected patient	ST5
BMD1232	HIV-uninfected patient	ST5
BMD1291	HIV-uninfected patient	ST5
BMD1353	HIV-uninfected patient	ST5
BMD1452	HIV-uninfected patient	ST5
BMD1592	HIV-uninfected patient	ST5
BMD1646	HIV-uninfected patient	ST5
BMD1716	HIV-uninfected patient	ST5
BMD1828	HIV-uninfected patient	ST5
BMD367	HIV-uninfected patient	ST5
BMD394	HIV-uninfected patient	ST5
BMD494	HIV-uninfected patient	ST5
BMD534	HIV-uninfected patient	ST5
BMD700	HIV-uninfected patient	ST5
BMD761	HIV-uninfected patient	ST5
BMD854	HIV-uninfected patient	ST5
BMD865	HIV-uninfected patient	ST5
BMD973	HIV-uninfected patient	ST5
LD1	Environmental	ST4
LD2	Environmental	ST5
NT7533	Environmental	ST5

Table 4-1 Clinical and environmental isolates of *C. neoformans* for comparison of pathogenicity according to lineage and source

4.3.2 *Galleria mellonella* infection model

The detailed procedure for maintaining and infecting larvae is presented in Appendix D. The procedure was based on the methods described by Firacative and Kavanagh *et al*[166, 229].

4.3.2.1 Preparation of larvae

Late instar isogenic wild type *Galleria mellonella* larvae (30 days from oviposition of the adult moth) were sourced from U U Pet Shop (Ho Chi Minh city, Vietnam). All larvae were kept in air porous containers at 16^{o}C in bran in the dark. Larvae for experimentation were selected according to the following criteria: weight 250-300 mg, healthy colour (beige). Grey larvae (indicating poor health) were discarded. All larvae were moved to room temperature for 2 hours prior to inoculation.

To control for potential variation in larval phenotype between batches, all survival experiments were internally controlled – that is I never used historical data regarding survival with a particular lineage. All infection experiments consisted of the strain of interested and simultaneous infection of Galleria from the same batch with the comparator strain of interest.

4.3.2.2 Identification of *Galleria mellonella larvae*

The larva species was confirmed morphologically. The greater wax moth larva (*G. mellonella*) is easily distinguished from the lesser wax moth larva (*Achroia grisella*) using a magnifier. The former have 4 stemmata on each side of the head while the latter do not have this characteristic.

4.3.2.3 Inoculum preparation

Yeasts were revived from frozen stock (Microbank Beads) on SDA plates for 72 hours. Single colonies were picked and cultured in YPD broth for 24 to 48 hours in a shaking incubator (SI-300, Jeio Tech), at 150 rpm at 30 ^{o}C until the concentration of cells was as at least 10^{8} cells/mL (titrated using a Cellometer X2 (Nexcelom Bioscience, USA). The culture was spun at 8000 rpm for 1 minute to form a pellet, subsequently washed twice with PBS. It was resuspended in PBS for quantitation

using the cellometer. I adjusted the suspension to a density of 10^8 cells/mL. Ten µL of inoculum (10^6 cells/larva) was used for injection per larva. All inocula were relabeled by an independent person such that I was blind to the nature of the inoculum used in each batch of larvae (i.e. the survival experiments were blinded).

4.3.2.4 Infection

Hamilton syringes were cleaned using 10% bleach by drawing and pushing out for 3 times followed by 100% ethanol and 1X PBS. The needle and syringe were cleaned between fungal strains. I used a cotton bud soaked in 70 % ethanol to swab the larva pro-legs prior to injection. Infection was performed by injecting 10 µL of inoculum (10^6 cells per larvae) into the last left pro-leg. Each cryptococcal strain was injected into 15 or 20 larvae. After injection, larvae were put into a petri dish covered with tissue for a few minutes to stop hemolymph leakage. I also injected 10 µL of PBS as control for sterility for every batch with the same number of larvae for each tested arm. Larvae infected with the same cryptococcal isolate were put into a petri dish and incubated at 37^{o}C for ten days. Larvae were checked twice daily for survival using physical stimulation with forceps. Larvae that did not respond to touch were considered to have died.

4.3.2.5 Blinding

To reduce any risk of bias in performing the experiment and assessing outcome, all inocula were relabeled by a third party such that I was unaware of which isolate was in use.

4.3.2.6 Controls

All experimental sets used the same batch of *Galleria*. PBS controls were included with all experimental sets.

4.3.2.7 Statistical analysis

Kaplan-Meyer survival curves were plotted using ggplot2 and survminer package (R version 3.4.2). The Log-rank test was used to compare differences in survival. Hazard Ratios were estimated using the Cox proportional-hazards model. Logistic regression

was used to investigate the change in the risk of larva death per every log10 increase in inoculum. The Wilcoxon rank sum test was used to compare fungal burden between isolates. Capsule size was compared using the Kruskal-Wallis test with P value adjusted for multiple comparisons by the Benjamini-Hochberg method.

4.3.3 Induction of *C. neoformans* var. *grubii* pathogenicity

Preparation of culture filtrate

Clinical isolates were grown to stationary phase (48 hours) followed by centrifugation at 8000 rpm for 1 min. The resulting supernatant was filtered using a Millipore membrane filter 0.45 μm (Merck), and was then checked for sterility by plating onto Sabouraud's agar and BHI broth and incubated at 30^{o}C for three days.

Preparation of inducing broth

Inducing broth consisted of 5 mL of fresh sterile YPD to which 2.5 mL of sterile culture filtrate was added (see above). Naive strains were inoculated into inducing broth and incubated for 48 hours in shaking incubator at 30^{o}C. Pellet was collected by centrifuging culture at 8000 rpm and washed with PBS twice prior to infection. I prepared an inoculum of 10^{8} cells/mL in PBS and 10 μL was injected into larvae. The workflow of inducing pathogenicity of naïve *C. neoformans* is described in the Figure 4-1.

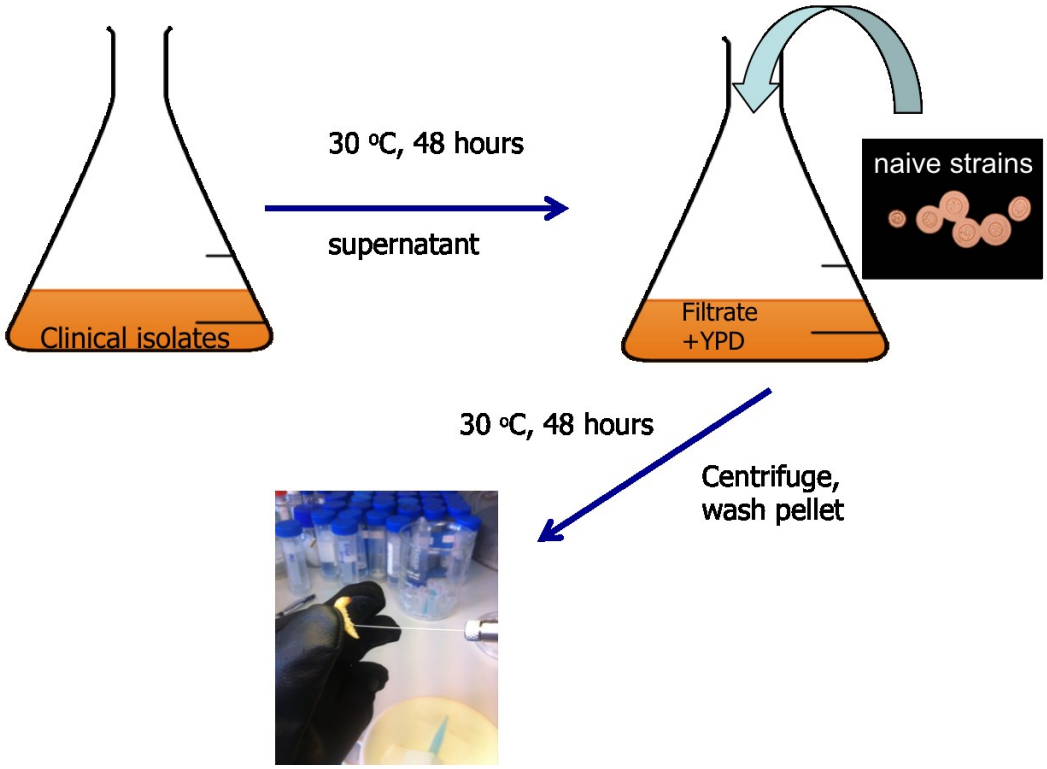


Figure 4-1 Inducing pathogenicity of naive environmental *C. neoformans*. Culture filtrate of clinical isolates was collected 48 hours post-inoculation. Naive environmental strains were then grown in fresh YPD plus the culture filtrate for another 48 hours. Pellet was washed and resuspended in PBS followed by adjustment to 10^8 cells/mL for larval infection.

4.3.4 Treatment of filtrate with freezing, heat, protease, and nuclease

To check the stability of culture filtrate exposed to prolonged frozen storage, the culture filtrate of BMD761 was kept at -20 ^{o}C for 33 days. The filtrate was brought to room temperature prior to induction experiments. Denaturation of filtrate was performed as described previously[230]. To denature protein by heat, filtrate of BMD761 was boiled at 97 ^{o}C for 2 minutes. To denature protein chemically, filtrate was incubated with Proteinase K (Sigma) at the concentration of 100 μg/mL in the presence of 30 mM Tris.HCl (pH=8) and 5 mM $CaCl_2$ at 37 ^{o}C for 1.5 hours. To degrade DNA, 28 μL DNAse I (Ambion) was added to 700 μL DNAse I buffer and 7 mL filtrate for 1 hour at 37^{o}C. To remove RNA, 2.5 mL filtrate was incubated with 50 μL RNAse cocktail of RNAse A and RNAse T1 (Thermo Fisher Scientific) for 1 hour at 37^{o}C.

4.3.5 Quantification of fungal burden in larvae

I collected dead larvae infected with strains of interest (5 dead larvae/biological replicate). Larvae were gently squeezed to collect hemolymph and fat body (about 1 gram) using a sterile knife. Hemolymph and fat body were collected into a 1.5 mL tube. I added sterile glass beads and 1mL of sterile water followed by homogenizing the mix at the frequency of 30 times/second for 10 minutes (TissueLyser II, Qiagen). I divided a Sabouraud agar plate into two equal sides then inoculated 100 µL of homogenate onto each side of the plate and incubated for 3 days at 30^{o}C. Finally I counted the number of colonies and calculated the CFU/gram/larvae.

4.3.6 Passage of *C. neoformans* through *G. mellonella* larvae

Following infection of *Galleria* as described previously yeast cells were harvested from the hemolymph of fresh (within 24 hours) dead larvae by inoculating hemolymph onto Sabouraud's agar plates and incubated at 30^{o}C for 3 days. Purified colonies were suspended in PBS and adjusted to 10^{8} cells/mL. 10 µL was reinfected into fresh *Galleria* larvae (20 larvae infected each generation) and incubated at 37 ^{o}C for 10 days. This cycle was repeated for 6 generations for each strain. Pathogenicity of the 6th generation *C. neoformans* and naive strain were compared using the larva killing assay as described above.

4.3.7 Capsule size measurement

Following infection with 10^{6} cells, larvae were sacrificed 48 hours post-infection and hemolymph extracted as in section 4.3.5 above. Hemolymph was stained with India ink to visualise capsule using light microscopy (at 1000X magnification). The diameter of the whole yeast cell and internal cell body were measured using ImageJ (https://imagej.nih.gov) based on the ruler bar embedded into each image. Capsule thickness was estimated by subtracting the internal cell diameter from the whole yeast cell diameter. Capsule size was calculated by subtracting the diameter of the cell body from the whole cell.

4.3.8 Experimental workflow

Experimental workflow for this chapter is depicted below

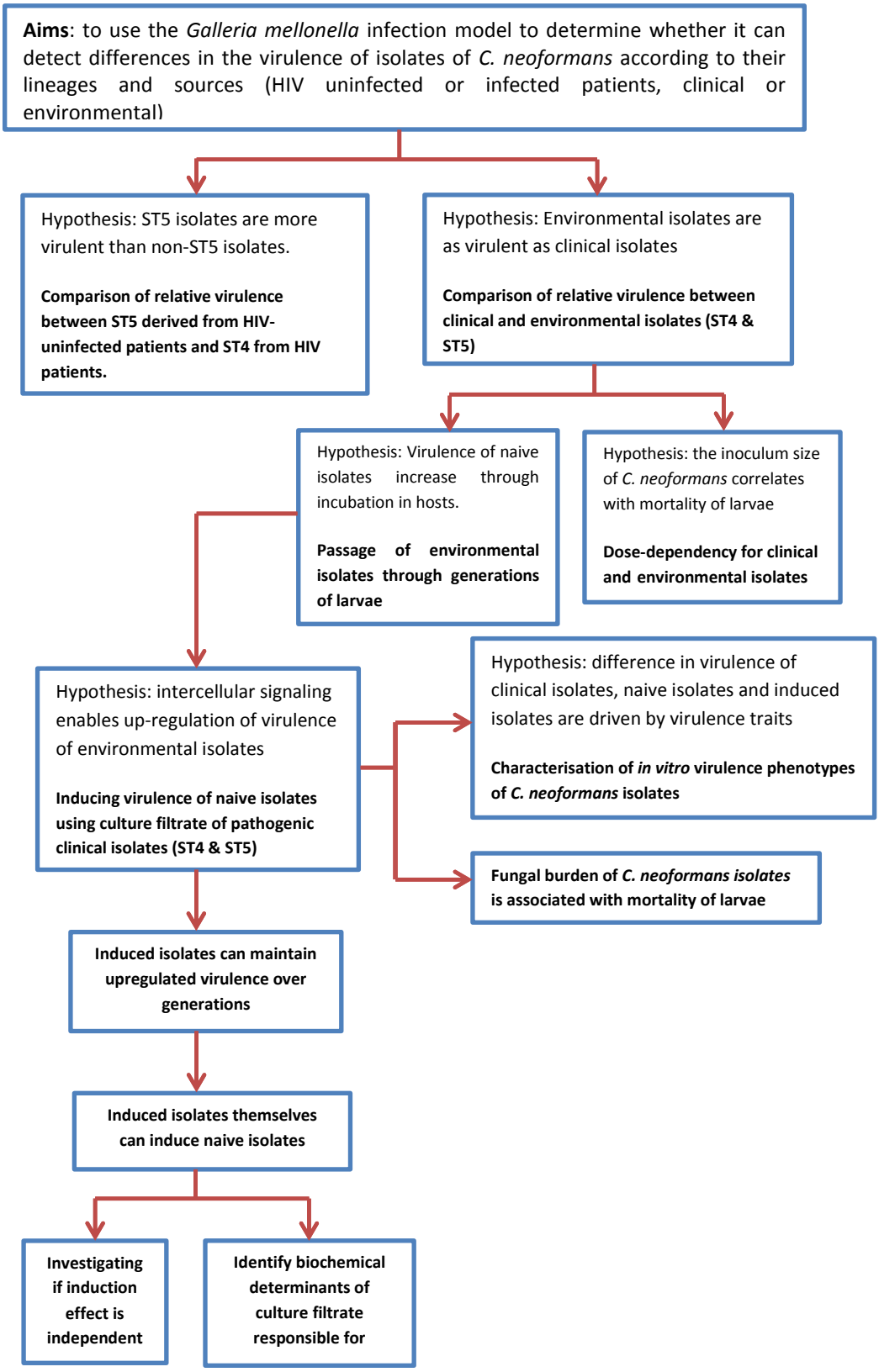


Figure 4-2 Experimental workflow for chapter 4

4.4 Results

4.4.1 ST5 isolates derived from HIV uninfected patients are more virulent than ST4 isolates derived from HIV infected patients

Each of the 20 ST4 or ST5 strains was inoculated into 15 larvae (N=600, 300 per arm). *C. neoformans* ST5 strains derived from HIV-uninfected patients were significantly more virulent than *C. neoformans* ST4 strains from HIV patients (HR 1.4, 95% CI 1.2 - 1.6, P value <0.001) (Figure 4-3).

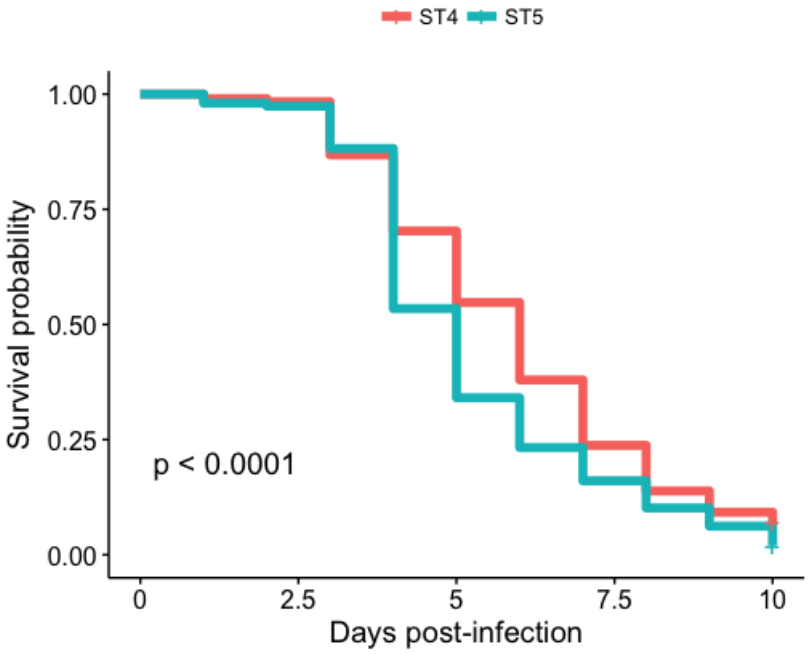


Figure 4-3 Survival curves of *G. mellonella* infected with *C. neoformans* strains: ST4 from HIV patients (20 isolates) vs ST5 from HIV-uninfected patients (20 isolates). P value was obtained from Log-rank test. The sample size is 600 larvae (15 larvae/strain)

4.4.2 ST4 and ST5 isolates derived from human patients are more virulent than strains derived from the environment.

Cryptococcosis is acquired from the environment and is not spread from person to person. Therefore, all strains are ultimately of environmental source and all strains of the same lineage would be expected to have the same ability to cause human disease. Moreover, ST5 strains derived from HIV-uninfected patient are interspersed in the whole-genome phylogeny of lineage 5, suggesting all ST5 strains have the

potential to cause disease in HIV-uninfected patients (Ashton P *et al*, unpublished data). Therefore I hypothesized that the pathogenicity of clinical and environmental isolates are indistinguishable, and tested this in the *Galleria* model. The isolates tested, and their sources and sequence types, are shown in Table 4-1. BK224 and LD1 are clinical and environmental isolates of ST4, respectively. BK42 plus BK139 and LD2 plus NT7533 are clinical and environmental isolates of ST5, respectively. I tested strains with MLST-matched genotypes.

The Kaplan Meier survival curves are shown in Figure 4-4. Clinically derived isolates were significantly more virulent than strains derived from the environment (HR 11.3, 95% CI 6.7-18.8, P value <0.0001). All *Galleria* larvae infected with clinically-derived strains died by day 8 regardless of sequence type, whereas a number of *Galleria* larvae infected with environmental isolates survived longer than 10 days.

I quantified the fungal burden in larvae infected with clinical isolates (BK139, BK224 and BK42) and those infected with environmental isolates (LD1, LD2 and NT7533). Fungal burden was significantly higher in the former (P value < 0.001, Wilcoxon test) (Figure 4-4).

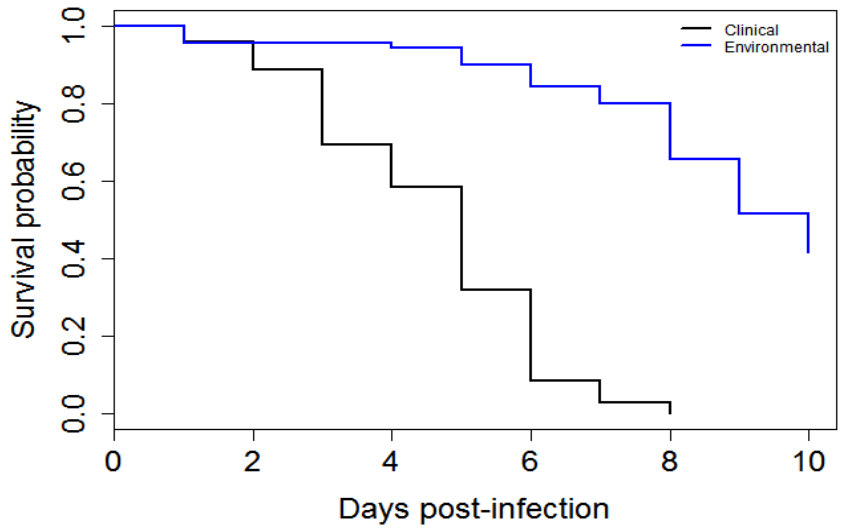


Figure 4-4 Survival analysis of clinical and environmental isolates. Six ST4 and ST5 isolates of clinical and environmental origin were infected with larva. Sample size is 72 larvae/arm. P value < 0.001, Log-rank test.

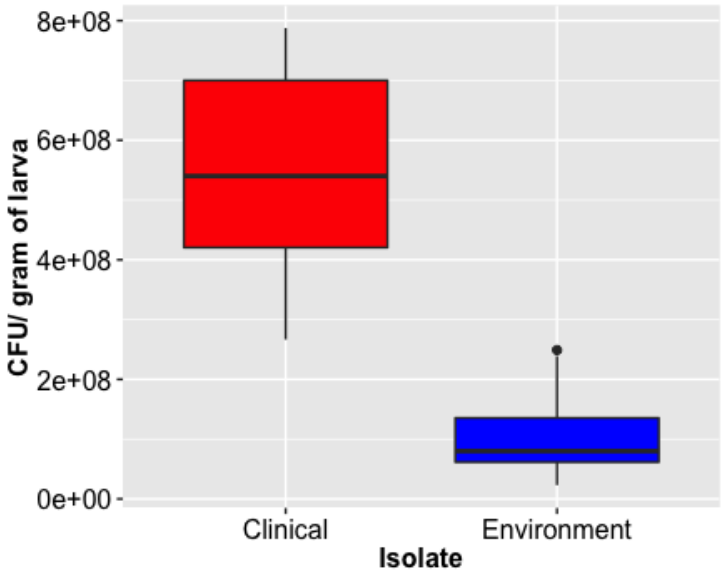


Figure 4-5 Fungal burden in larvae infected with clinical (BK224, BK42, BK139) and environmental isolates (LD1, LD2 and NT7533). P < 0.001 (Wilcoxon test). Sample size is 24. Each strain was quantified for fungal burden with 4 biological replicates.

4.4.3 Dose-dependent pathogenicity

I hypothesized that size of the inoculum of *C. neoformans* correlates with mortality of larvae. I chose two strains of the same sequence type 5 but derived from different sources: environmental LD2 and clinical BMD761. Each strain was infected into larvae with an inoculum varying from 10^3 to 10^6 (40 larvae/inoculum size) and larval survival was observed for 10 days at 37^{o}C. I found that for every log10-cells increase in inoculum size of the LD2 isolate, the odds of larval death increased by 3.4 (OR 3.4, 95% CI 2.3-5.4, P value < 0.001). For the clinical BMD761 isolate, for every log10 increase in inoculum size, the odds of larval death increased by 7.3 (OR 7.3, 95% CI 4.3-14.1, P value < 0.001). Predicted probabilities of larval death are illustrated in the figure 4-6 and 4-7 below.

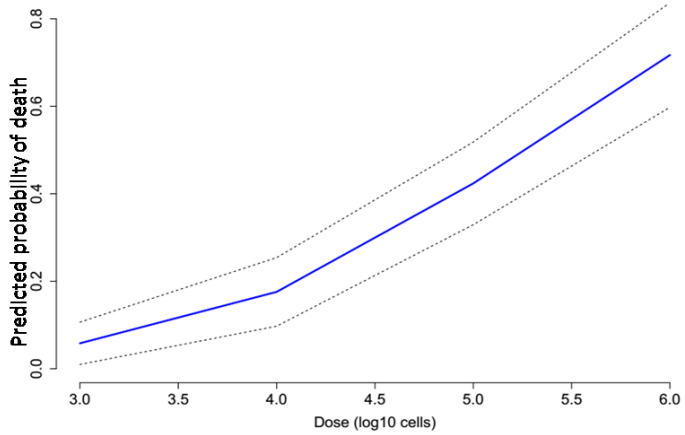


Figure 4-6 Logistic regression model of the risk of death in relation to inoculum size for the environmental *C. neoformans* isolate LD2. Predicted probabilities and 95% CI are blue line and dot line, respectively. Doses vary from 10^3 to 10^6 cells/inoculum. X axis indicates inoculum (log10 cells). Y axis indicates probability of death of larvae infected with LD2.

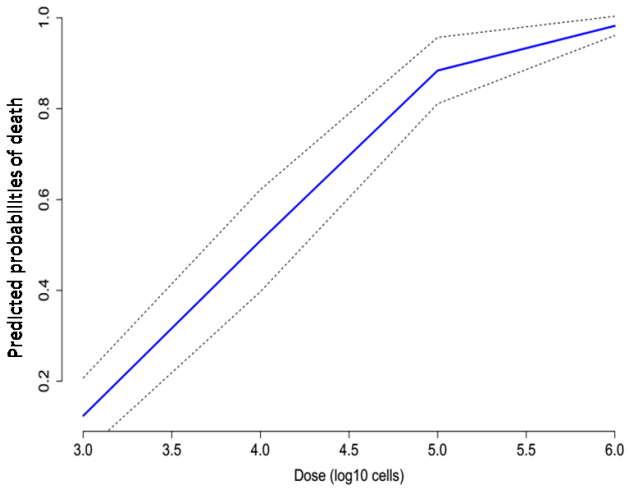


Figure 4-7 Logistic regression model of the risk of death in relation to inoculum size for the clinical *C. neoformans* isolate BMD761. Predicted probabilities and 95% CI are shown by the blue and dotted lines, respectively. Doses vary from 10^3 to 10^6 cells/inoculum. X axis indicates inoculum (log10 cells). Y axis indicates probability of death of larvae infected with BMD761.

4.4.4 Virulence of environmental *C. neoformans* is fixed following passage through the larvae.

Cryptococcosis in most cases is considered to occur as the result of reactivation of latent infection. Latent infection follows an initial infection that is asymptomatic in the vast majority of cases[72]. I hypothesized that *C. neoformans* 'learns' increased virulence through adaptation to the host following infection, as a result of low level growth. To determine whether virulence can be upregulated within the host I repeatedly passaged environmental strains in *Galleria* larvae and compared virulence with the baseline status of the strain. Naive environmental *C. neoformans* strains (LD1 and NT7533 representing ST4 and ST5, respectively) were passed through 6 generations of larvae and then subjected them to the *G. mellonella* killing assay. However, I found no change in virulence between the index and F6 strains (Figure 4-8).

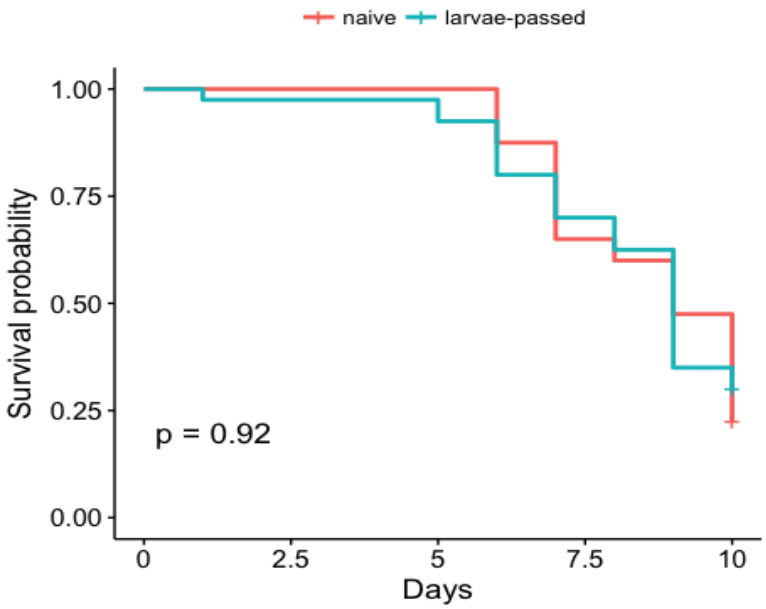


Figure 4-8 Survival curves of *G. mellonella* infected with naive vs larvae-passed environmental isolates. Sample size is 80 (40 larvae/arm). P value was obtained from Log-rank test.

4.4.5 Pathogenicity induction of environmental *C. neoformans var. grubii* isolate

In the following experiments, I tested if pooled filtrate of ST5 isolates derived from HIV-patients (BK139 and BK142 isolates) or non-HIV patients (BMD854, BMD761,

BMD494 and BMD973 isolates) can induced ST5 environmental isolates (LD2 and NT7533). Then I investigated if individual ST5 clinical isolates above can induce naive environmental isolates.

4.4.5.1 Isolates derived from HIV uninfected patients can upregulate the virulence of environmental isolates of *C. neoformans* ST5.

AFLP data previously published from our lab suggest that strains derived from HIV uninfected apparently immunocompetent patients are spread throughout the entire VNIg (ST5) lineage. This implies that all strains have the potential to infect immunocompetent patients, despite the differences I saw between environmentally and clinically derived strains, and that virulence may be up (or down) regulable. Presumably all cells of a particular isolate must behave similarly in order to deliver a specific virulence phenotype; maintenance of this phenotype amongst cells may be driven by intercellular signaling. To test this hypothesis that intercellular signaling may enable up-regulation of virulence, I cultured naive environmental isolates (NEI) ST5 strains in pooled culture filtrate from clinical isolates (BK139 and BK42 representing ST5) derived from HIV patients. The resulting strains were injected into larvae in the usual manner. There was a separation of the survival curves for these 'induced' and NEI isolates, but the effect was modest and did not reach conventional levels of statistical significance (Hazard ratio 1.4, 95% CI 0.8-2.3, P=0.2) (Figure 8A). I repeated the experiment but used culture filtrate from ST5 strains derived from HIV-uninfected patients (BMD854, BMD761, BMD494 and BMD973). This resulted in a highly significant increase in *Galleria* mortality (Hazard ratio 4.1, 95% CI 2.9-5.8, P value <0.001) (Figure 8B).

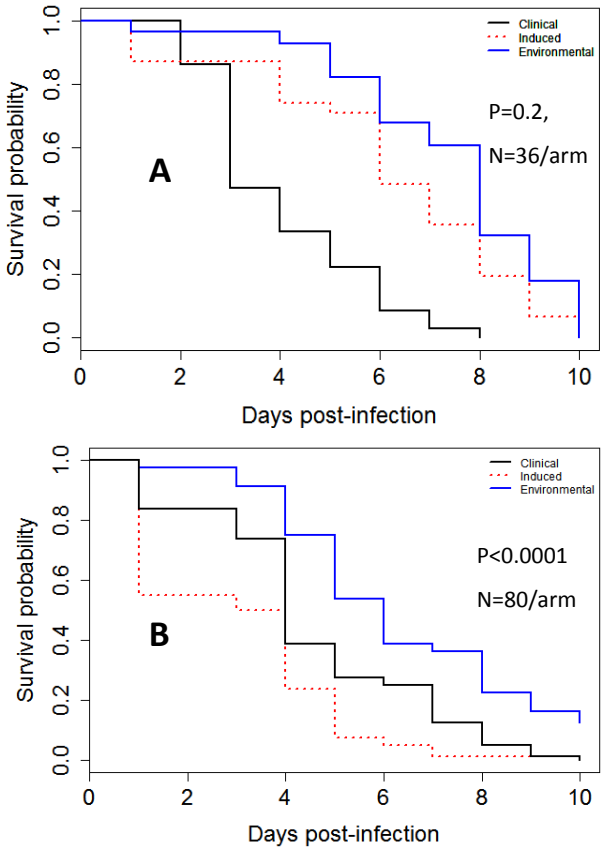
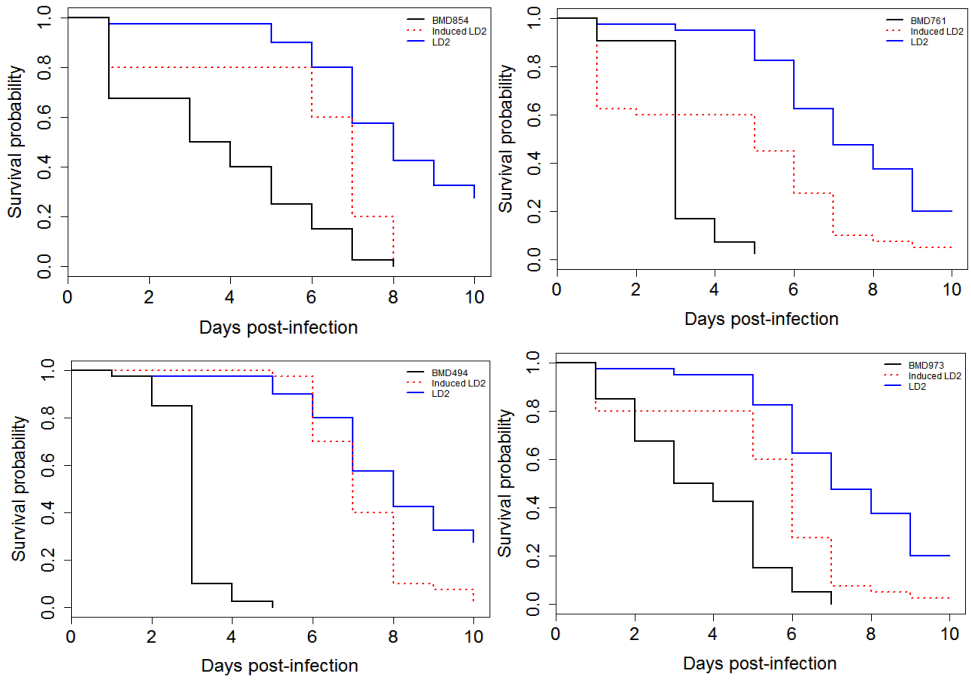


Figure 4-9 Survival curves of larvae infected with clinical, naive environmental and induced environmental isolates. Environmental isolates (LD2 and NT7533) were induced using pooled culture filtrate from ST5 strains (n=2) from HIV patients (**A**) or ST5 isolates (n=4) from HIV uninfected patients (**B**). P values were obtained from Cox proportional-hazards model for comparison between induced and naive isolates.

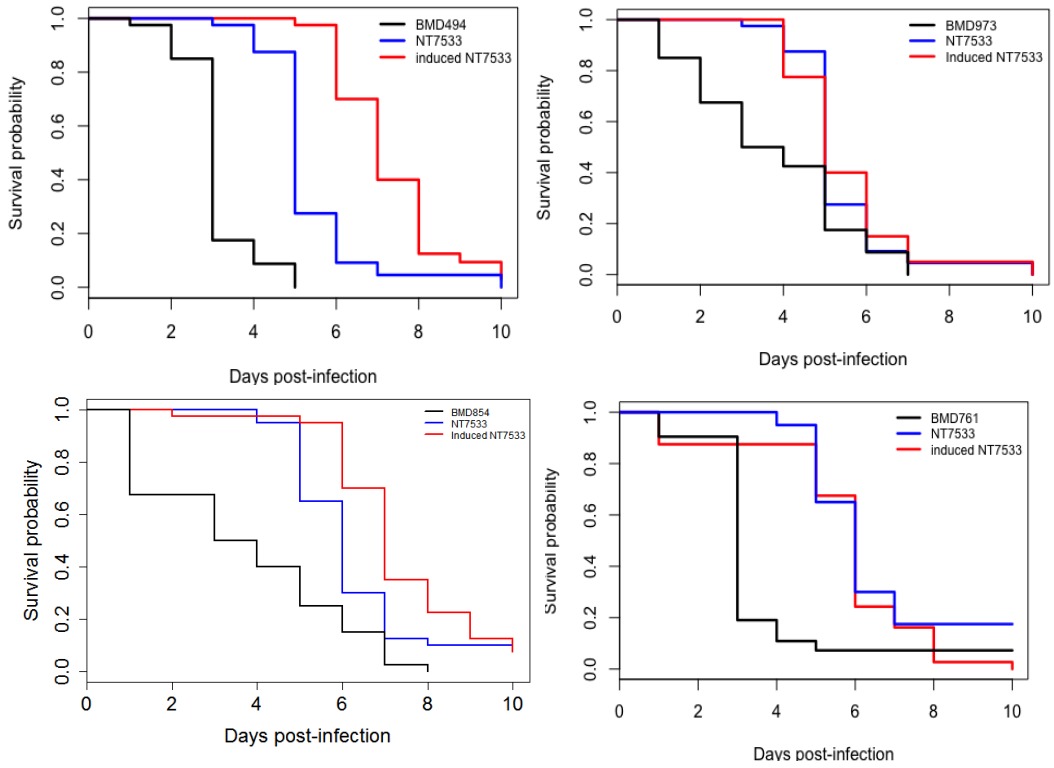
4.4.5.2 All ST5 strains derived from HIV-uninfected patients can induced the naive environmental strain LD2

To test the individual 'inducing' effect of each clinical strain, I used individual culture filtrates from BMD854, BMD761, BMD494 and BMD973 to induce environmental isolates LD2 (ST5) and NT7533 (ST5) individually. All four strains exhibited the ability to upregulate the virulence of LD2 (Figure 9). In contrast none of NT7533 strains were induced. Surprisingly, NT7533 strains grown in culture filtrate from BMD494 or BMD854 were significantly less pathogenic than naive NT7533 (Figure 10).



Pairwise comparison	Hazard ratio	95% CI	P value
BMD494-induced LD2 vs LD2	2.2	1.3-3.6	0.002
BMD973-induced LD2 vs LD2	**2.7**	**1.7-4.4**	**<0.001**
BMD854-induced LD2 vs LD2	3.4	1.9-5.8	<0.001
BMD761-induced LD2 vs LD2	**2.7**	**1.7-4.8**	**<0.001**

Figure 4-10 Survival curves of larvae infected with clinical, naive environmental and induced environmental isolates. Environmental isolates (LD2) were induced by individual culture filtrates of clinical BMD854, BMD761, BMD494 and BMD973 isolates. P values and hazard ratios obtained from Cox proportional-hazard model, were applied for induced isolates and naive ones. Sample size is 40 larvae/arm



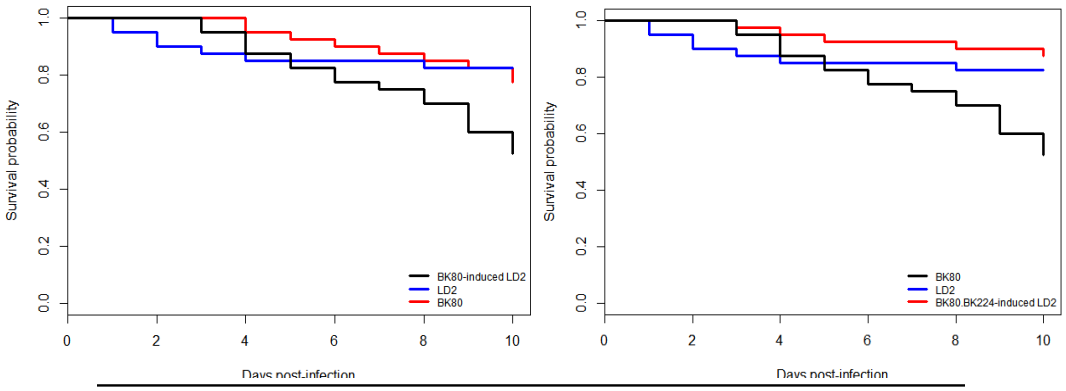
Pairwise comparison	Hazard ratio	95% CI	P value
BMD494-induced NT7533 vs NT7533	0.2	0.1-0.4	<0.001
BMD973-induced NT7533 vs NT7533	**1.1**	**0.6-1.4**	**0.6**
BMD854-induced NT7533 vs NT7533	0.6	0.4-1	0.03
BMD761-induced NT7533 vs NT7533	**0.8**	**0.5-1.2**	**0.26**

Figure 4-11 Survival curves of larvae infected with clinical, naïve environmental and induced environmental isolates. Environmental isolates (NT7533) were induced by individual culture filtrates of clinical BMD854, BMD761, BMD973 and BMD494 isolates. P values and hazard ratios obtained from Cox proportional-hazard model, were applied for induced isolates and naive ones. Sample size is 40 larvae/arm.

4.4.5.3 ST4 strain-derived culture filtrate is unable to upregulate the pathogenicity of ST5 environmental strains

To test the ability of clinically derived ST4 strains to upregulate the pathogenicity of environmental strains, I used the culture filtrate of BK80 strains to grow LD2 strains as described previously. The virulence of naive LD2 and ST4-induced LD2 was not significantly different (HR 0.8, 95% CI 0.3-2.2, P value=0.7) (Figure 4-12). I repeated induction of LD2 using pooled supernatant from both BK80 and BK224. However,

there was still no effect. The survival rate of larvae infected with induced LD2 was not significantly different from those infected with naive LD2 (HR 1.5, 95% CI 0.5-4.8, P value=0.5). ST4 strains are necessarily from HIV infected patients since this lineage does not infect apparently immunocompetent patients. Therefore I could not perform this experiment with culture filtrate derived from ST4 strains from apparently immunocompetent patients.

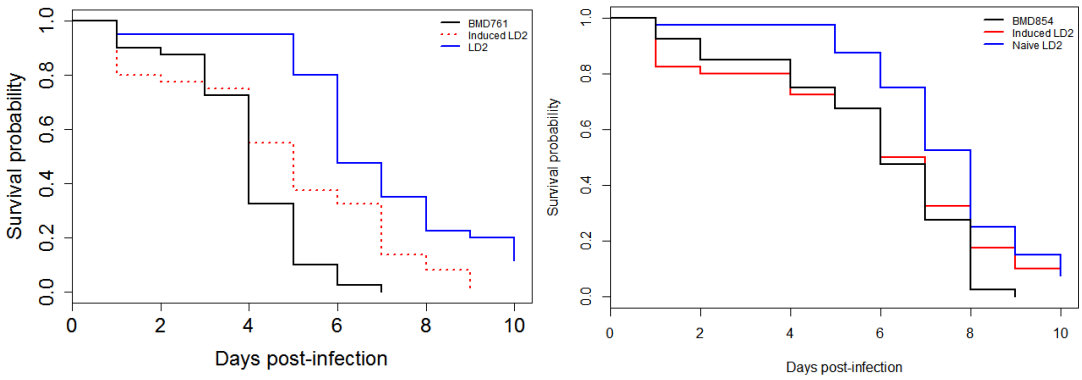


Pairwise comparison	Hazard ratio	95% CI	P value
LD2 vs BK80-induced LD2	0.8	0.3-2.2	0.7
LD2 vs BK80.BK224-induced LD2	**1.5**	**0.5-4.8**	**0.5**

Figure 4-12 Survival curves of larvae infected with clinical, naïve environmental and induced environmental isolates. LD2, environmental isolate, was induced by individual filtrate (BK80) or pooled filtrate (BK224 + BK80). Sample size is 40 larvae/arm

4.4.6 Upregulated virulence is persistent over generations

To test if induced isolates maintain their up-regulated virulence over one generation, I extracted induced isolates from larvae, cultured on Sabouraud agar and then repeated the infection experiments. The BMD761-induced LD2 strain clearly maintained its upregulated virulence; however, LD2 induced by BMD854 appeared to lose some virulence, although the survival curves remained separated (Figure 4-13). Hence the BMD761 strain is used for downstream experiments. The ability of BMD761 to induce LD2 was confirmed by repeating the experiment on 6 independent occasions. The results was presented in the appendix E.



Pairwise comparison	Hazard ratio	95% CI	P value
LD2 vs BMD854-induced LD2	0.7	0.4-1.1	0.1
LD2 vs BMD761-induced LD2	**0.4**	**0.3-0.7**	**<0.001**

Figure 4-13 Survival curves of clinical isolates, naive environmental isolates and induced environmental isolates revived from larvae infected with induced isolates. BMD761 and BMD854 are clinical isolates while LD2 is derived from the environment. N= 40 larvae /arm

4.4.7 Induced isolates can induce naive isolates

I next tested whether the induced LD2 isolate could transform its naive self. I grew the naive LD2 in culture filtrate of a previously BMD761-induced LD2 strain. The virulence of the LD2-induced LD2 strain significant increased compared to naive LD2 (Hazard ratio 2.8, 95% CI 1.7-4.4, P value <0.001) (see Figure 4-14)

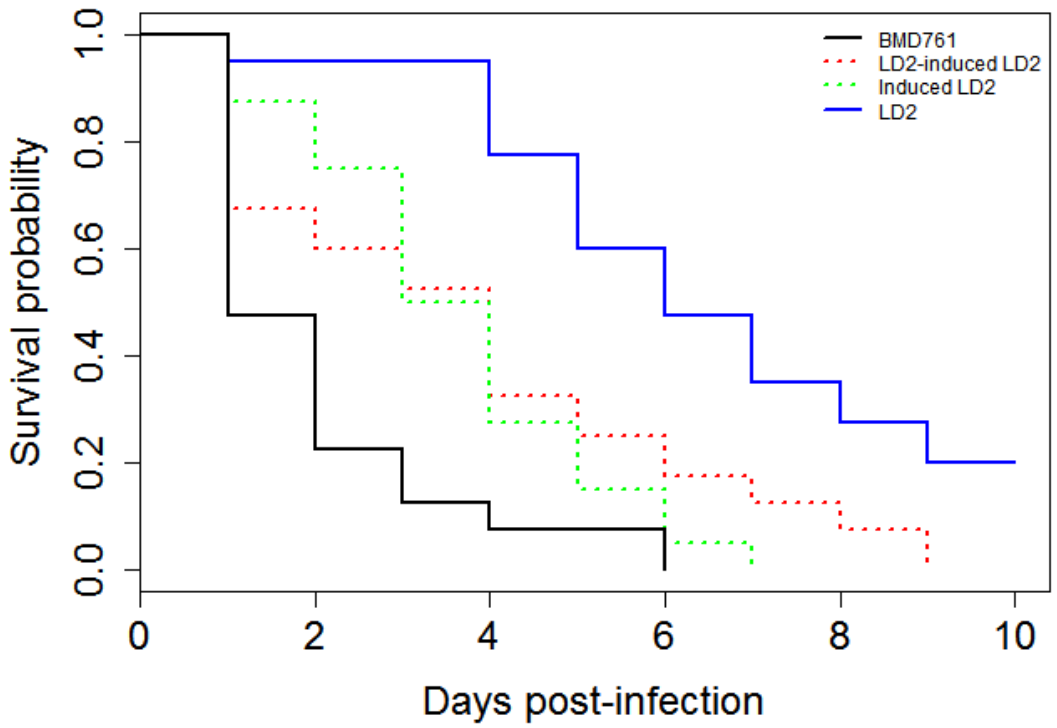


Figure 4-14 Survival curves of clinical (BMD761), environmental (LD2), induced isolates (Induced LD2 and LD2-induced LD2). Induced LD2 indicates BMD761-induced LD2. LD2-induced LD2 indicates naive LD2 transformed by BMD761-induced LD2. P value for comparison of L2-induced LD2 vs naive LD2 was obtained from the Cox proportional-hazards model.

4.4.8 Virulence induction is not a consequence of growth in nutrient depleted media

To test whether virulence induction was an artefact of growing a strain in a somewhat nutrient depleted medium, I cultured the naive isolate in nutrient-deplete culture filtrate. Here, instead of culturing naive LD2 in culture filtrate from BMD761, I grew it in culture filtrate from naive LD2. Nutrient-deplete culture filtrate was obtained by inoculating LD2 strains into YPD broth for 48 hours followed by centrifuge and filtration (i.e the same method as used for the induction experiments described above). The virulence of self-induced LD2 strains was not significantly different to the naive strain (HR 1.2, 95% CI 0.8-1.9, P value=0.3) (Figure 4-15).

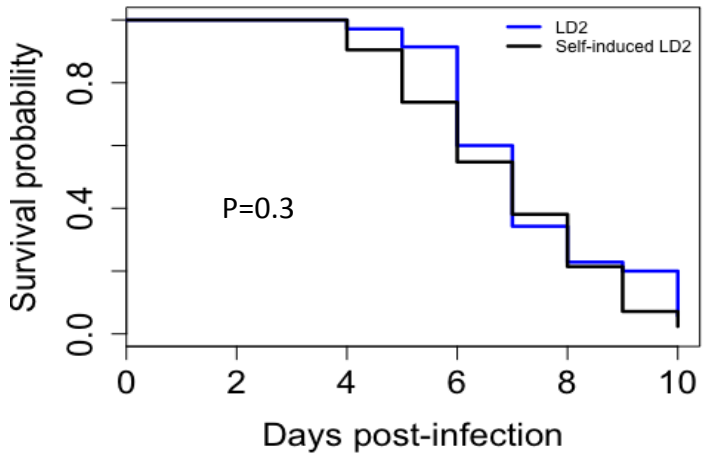
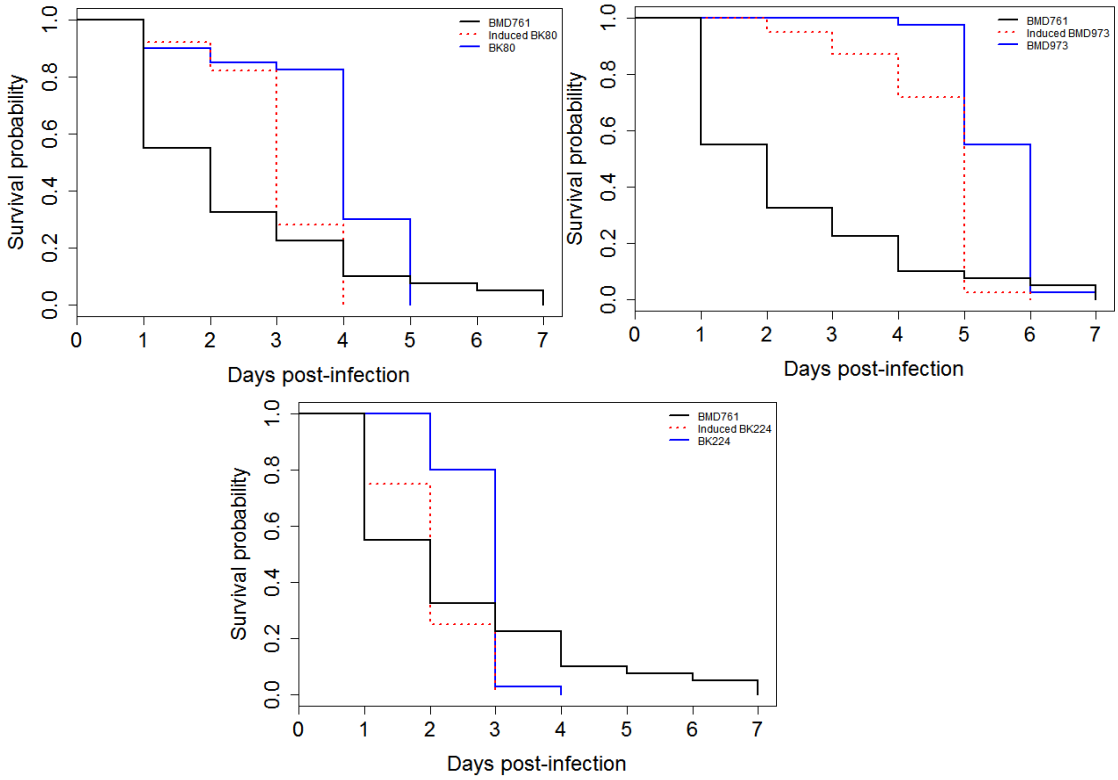


Figure 4-15 Survival curves of larvae infected with naive LD2 and self-induced LD2 strains. P value was generated using the log-rank test.

4.4.9 Hypervirulent isolates can induce upregulation of virulence of other clinical strains

I tested whether BMD761 could upregulate the virulence of other clinical strains. Using the methods already described I added BMD761 culture filtrate to broth cultures of strains BMD973 (ST5), BK80 & BK224 (ST4). Larvae infected with BMD761-induced strains exhibited worse survival than naive clinical strains (Figure 4-16). While the baseline virulence of BK224 was similar to that of BMD761 (HR 1.2, 95% CI 0.7-1.9, P=0.5), BMD761-induced BK224 was even more pathogenic than BMD761 (P=0.02).



Pairwise comparison	Hazard ratio	95% CI	P value
BMD761-induced BK80 vs BK80	2.2	1.4-3.5	<0.001
BMD761-induced BK224 vs BK224	**2.1**	**1.4-3.4**	**0.001**
BMD761-induced BMD973 vs BK973	2.2	1.4-3.6	<0.001

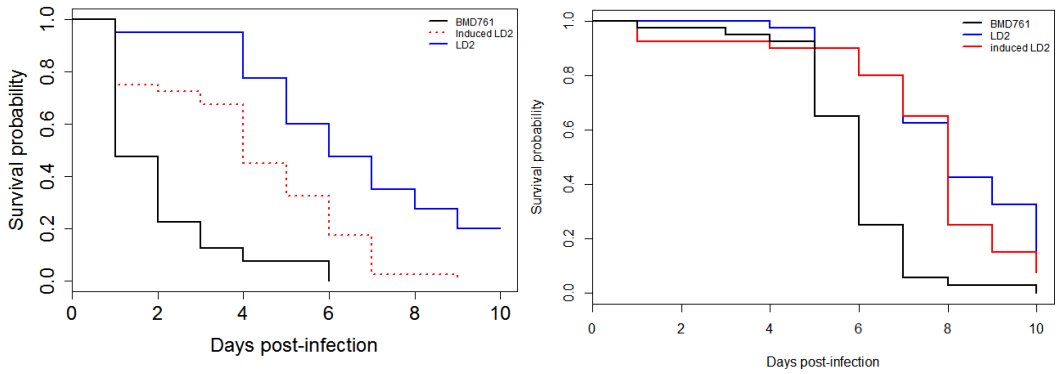
Figure 4-16 Survival curves of BK80, BK224 and BMD973 and corresponding BMD761-induced isolates. P values are derived from Cox's proportional-hazards model for comparison between induced isolates and naive ones**.** Sample size was 40 larvae/arm

4.4.10 The induction capacity of culture filtrate is lost by boiling and protease treatment

To test the stability of BMD761 culture filtrate activity under freezing for prolonged storage, I froze culture filtrate at -20^{o}C for 33 days. Frozen filtrate maintained its induction effect (Figure 4-17)

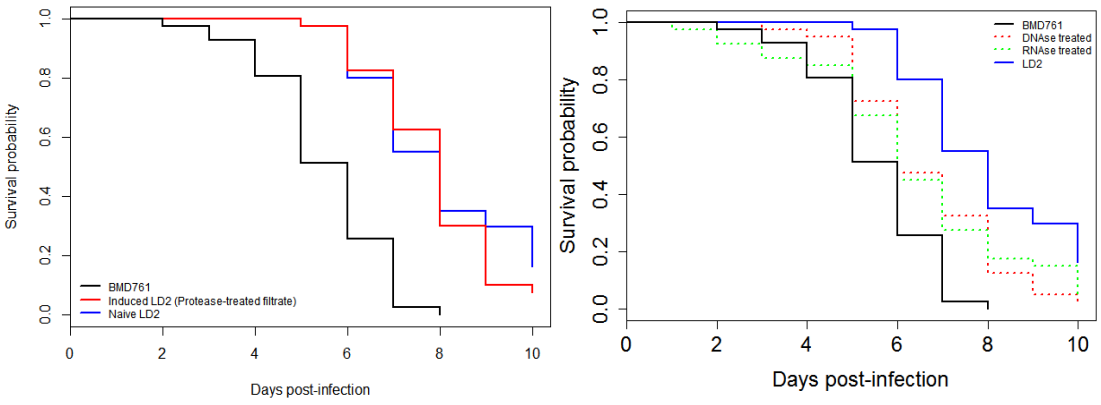
To understand whether nucleic acids or protein/peptide mediated the induction effect, BMD761 culture filtrate was incubated with RNAse, DNAse and protease

(proteinase K) and heat prior to induction. The results show that culture filtrate lost its effect when exposed to heat (Figure 4-17) or protease treatment (Figure 4-18).



Pairwise comparison	Hazard ratio	95% CI	P value
LD2 vs Induced LD2 (frozen filtrate)	0.4	0.2-0.6	<0.001
LD2 vs Induced LD2 (boiled filtrate)	**1.3**	**0.8-2.1**	**0.3**

Figure 4-17 Survival curves of larvae infected with clinical (BMD761), environmental (LD2) and induced isolated. LD2 was induced with either frozen (left plot) or boiled (right plot) culture filtrate. P values and HR were calculated using Cox proportional-hazards model.



Pairwise comparison	Hazard ratio	95% CI	P value
LD2 vs Induced LD2 (protease)	1.2	0.8-2	0.36
LD2 vs Induced LD2 (DNAse)	**2.1**	**1.3-3.4**	**0.002**
LD2 vs Induced LD2 (RNAse)	2	1.2-3.1	0.004

Figure 4-18 Survival curves of larvae infected with clinical (BMD761), environmental (LD2) and induced isolated. Culture filtrates were treated with nuclease or protease prior to induction. P values and HR obtained from the Cox proportional-hazards model are presented in the table above

4.4.11 Upregulated virulence is associated with larger fungal burdens in infected *Galleria*

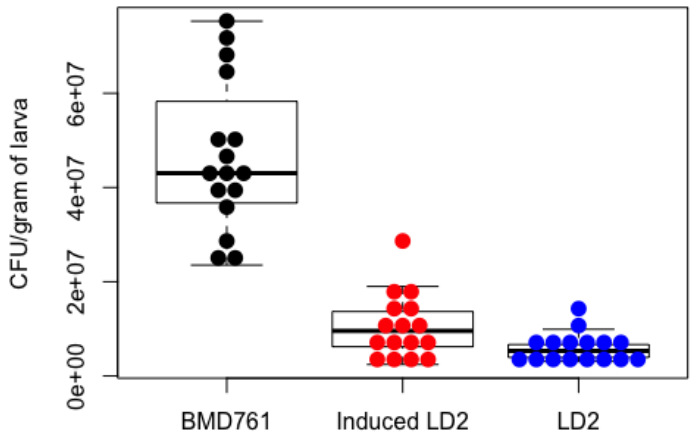


Figure 4-19 Fungal burden of *C. neoformans* isolates (clinical BMD761, environmental LD2 and induced LD2) recovered from larval hemolymph 48 hours post-inoculation. P value (Wilcoxon test) is 0.025 for comparison between LD2 and Induced LD2. N=48, 16 biological replicates/strain.

After inoculation for 48 hours, larvae were sacrificed to extract hemolymph and fat body which were cultured to quantify CFU/gram of larva. Fungal burden was significantly higher in larvae infected with induced LD2 compared with those infected with naive LD2 (P=0.025). The fungal burden in BMD761-infected larvae was more variable than that of LD2 and Induced LD2 (Figure 4-19).

4.4.12 Virulence upregulation is not associated with measurable differences in classical virulence traits

Thermotolerance, capsule size, melanin formation and CSF *ex vivo* survival are phenotypes previously associated with virulence in *C. neoformans* as mentioned in the introduction chapter. I tested whether larval virulence was associated with changes in phenotype by comparing phenotypes of induced LD2 with naive LD2. The methods of phenotypic assays were described in the chapter 3. There were no differences between induced and naive isolates across any of these parameters (P values for difference in growth at 30^{o}C and 37^{o}C were 0.2 and 0.4, respectively). CSF *ex vivo* survival testing for 72 hours at 37^{o}C was indistinguishable (P value =1). Similarly, no difference was seen in melanin production by visual observation using

Candida parapsilosis as control for negative melanization. Contrary to my expectation, the capsule size of naive LD2 isolates (extracted from larval hemolymph) was significantly thicker than induced LD2 (Figure 4-20). Both capsule size of LD2 and Induced LD2 were significantly smaller than that of BMD761 (Adjusted P value = 0.03 and < 0.001, respectively)

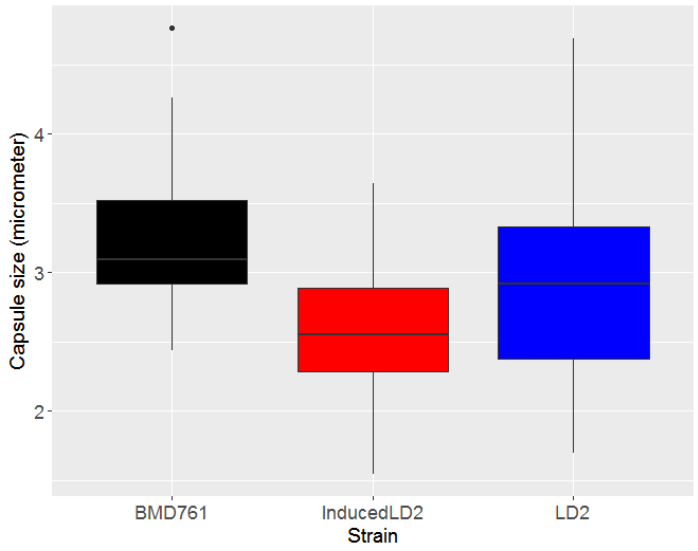


Figure 4-20 Capsular size comparison of *C. neoformans* (clinical BMD761, environmental LD2 and Induced LD2) recovered from larval hemolymph 48 hours post-inoculation. Adjusted P value = 0.03 obtained from Kruskal-Wallis test with Benjamini-Hochberg method (Induced LD2 vs LD2)

4.4.12 Morphology of *C. neoformans* post-inoculation in larvae

C. neoformans were extracted from hemolymph of living larvae 48 hours post-inoculation. All cells were highly encapsulated with pseudophyphae cells detected in induced LD2 strains (Figure 4-21). However, pseudophyphae cells were not frequently observed. Titan cells were observed in both LD2 and induced LD2 strains but not in BMD761 strains 96 hours post-inoculation. Titan cells are classified as those with cell body diameter greater 15 μm[128]. Titan cells were first described in the mouse lung. They are resistant to phagocytosis and oxidative and nitrosative stresses, and are believed to enhance cryptococcal survival *in vivo*[126 125].

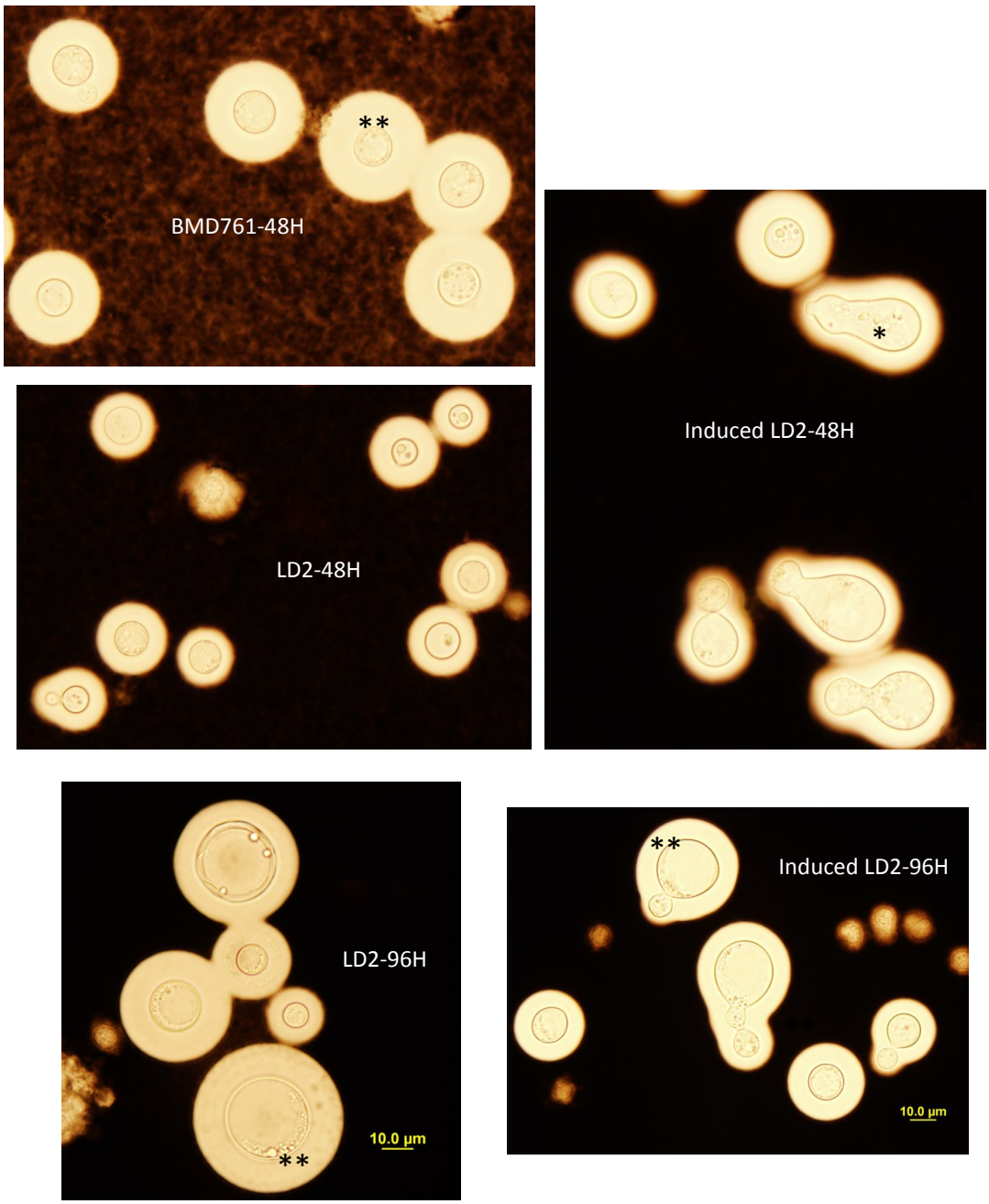


Figure 4-21 Morphology of C*. neoformans* extracted from larva 48 hours post inoculation at 1000X magnification under light microscope: clinical BMD761, naïve LD2 and Induced LD2, and 96 hours post-inoculation: naïve LD2 and Induced LD2. Titan cells and pseudohyphae cells are denoted with ** and *, respectively.

4.4.13 ST5 strains derived from HIV-uninfected patients appear more virulent than Vietnamese *C. gattii* strains in the *Galleria* model

Disease in immunocompetent people due to infection with *C. gattii* is well-described, and a small number of cases (2-6) are seen each year in our hospital. It is rare to

encounter a *C. gattii* strain in HIV infected patients in Vietnam. Given the fact that almost all disease in HIV uninfected patients due to *C. neoformans* in Vietnam is caused by ST5 strains, I was interested in whether the ST5 strain has comparable virulence to Vietnamese *C. gattii* isolates. I compared survival in the *Galleria* model using three *C. neoformans* ST5 strains (BMD1592, BMD367, BMD494) vs three *C. gattii* strains (BMD873, BMD1377, BMD800) - all derived from HIV-negative patients. The survival rate of larvae infected with *C. neoformans* ST5 strains derived from HIV negative patients was significantly lower than *C. gattii* (P value <0.0001, Figure 4-22).

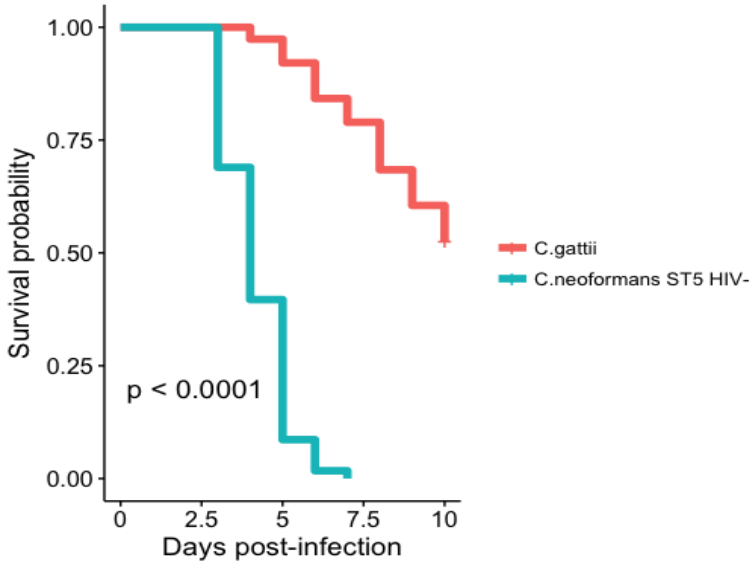
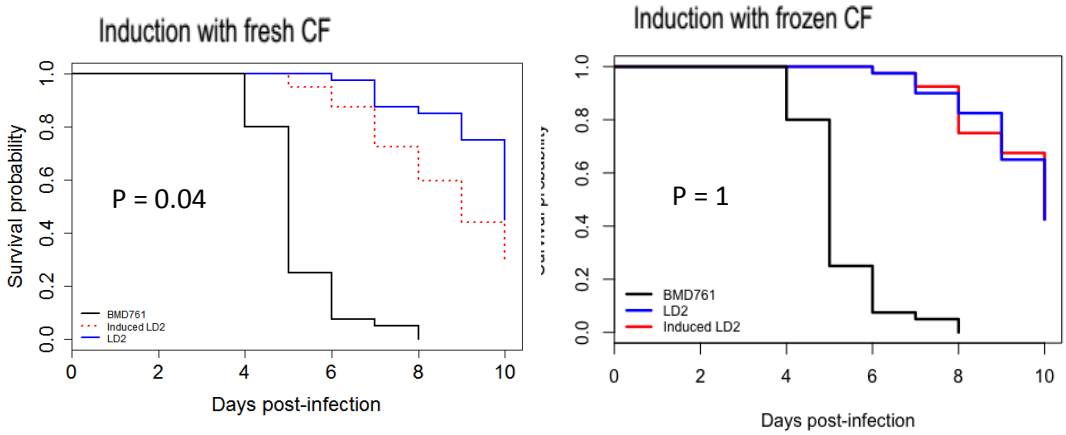


Figure 4-22 Survival curves of larvae infected with *C. gattii* and C*. neoformans* ST5 derived from HIV-uninfected patients. P value < 0 obtained from Log-rank test, n=90

4.4.14 Medium-dependent induction effect

I hypothesize that the induction effect is independent of medium. I employed RPMI instead of YPD to culture BMD761 and used the culture filtrate, both fresh and frozen, to induce LD2 strain. The induction effect was maintained when using RPMI, but only when fresh (as opposed to frozen) culture filtrate was used (Figure 4-23). However, the induction effect was not as robust as in the case of YPD-based culture.



Pairwise comparison	Hazard ratio	95% CI	P value
LD2 vs induced LD2 (Fresh CF)	0.55	0.3-0.96	0.04
LD2 vs induced LD2 (Frozen CF)	1	0.6-1.8	0.9

Figure 4-23 Survival analysis of larvae infected with clinical BMD761, naïve LD2 and LD2 induced with fresh and frozen RPMI-base culture filtrate. P value and HR were obtained from pairwise comparison between naïve LD2 and induced LD2 using Cox proportional-hazards model. N= 120

4.5 Discussion

4.5.1 Comparative survival of ST4 strains derived from HIV-infected and ST5 strains derived from HIV-uninfected patients

The mouse is an established model for studying pathogenicity of *C. neoformans* because it shares some similarity with humans, with regard to general anatomy, physiological temperature and immune system. However, it has the disadvantage of being expensive, low throughput, and sentient. Moreover, with regard to cryptococcosis, it may represent more a model of disease in HIV infection rather than immunocompetent infection [231].

In recent years, there has been increased interest in the *G. mellonella* infection model because of its relative ease of use, reduced bioethical issues, lower costs and potential for higher throughput. Despite being an invertebrate, *Galleria* has an immune system that shares considerable similarities with the mammalian innate immune system. *Galleria* hemocytes phagocytose bacterial and fungal cells in a

similar way to human neutrophils through the oxidative burst pathway. Proteins mediating generation of reactive oxygen species in *Galleria* larvae are homologous to proteins involved in superoxide production by human neutrophils [232]. Following these discoveries, *G. mellonella* has been widely adopted to study the pathogenesis of bacteria (*Bacillus cereus, B. thuringiensis & E. coli*) and fungi (*Candida albicans, C. krusei, Aspergillus fumigatus, A. flavus & Cryptococcus gattii*)[229] [166]. The *G. mellonella* infection model was previously used to distinguish the hyperpathogenic *C. gattii* strain causing Vancouver Island outbreak (VGIIa) and other major molecular types (VGII, VGIII and VGIV)[166]. Moreover, correlation between virulence of *C. albicans* and *Pseudomonas aeruginosa* in mice and larvae has been established[164] [163]. This supports the use of *Galleria mellonella* as a model of innate immunity for high-throughput pathogenicity assays instead of mammalian models.

I hypothesized that ST5 isolates are more virulent than non-ST5 isolates. Here, I found that *C. neoformans* ST5 isolates derived from HIV uninfected patients are significantly more virulent than non-ST5 strains (which are necessarily derived from HIV infected patients). I believe that enhanced virulence, when detected in a model of the innate immune system (ie. the *Galleria* model), is consistent with my hypothesis that ST5 strains have increased pathogenicity, where pathogenicity means the ability to cause infection/disease[86]. The innate immune response is the first line of barrier against cryptococcal infection with two major players – complement and phagocytic effector cells. The complement system is responsible for opsonising cryptococcal cells and attracting phagocytic effector cells. Phagocytosis is performed by macrophages and neutrophils in humans. Neutrophils are an integral component of the innate immune response to cryptococci [233]. They kill pathogens by exerting the oxidative burst and producing antimicrobial peptides (e.g. defensins).

4.5.2 Clinical *C. neoformans* are significantly more virulent than environmental isolates

I hypothesized that environmental isolates are as virulent as clinical isolates. The difference I identified in the pathogenicity of clinical isolates compared with environmental isolates of the same MLST is consistent with observations that clinical

isolates are more virulent than environmental isolates in the mouse model [52] [62]. Ngamkulrungroj *et al* observed a similar effect *with C. gattii* strains (VGIIb) [234]. The increased virulence of clinical strains compared with environmental strains could be explained by within-host microevolution[235], driven by selective pressures that are absent in the environment outside of the animal host. Infection results from the inhalation of (possibly dormant) yeasts or spores. During the prolonged incubation in human hosts, there may be low level yeast growth that stimulates a conversion to a raised virulence state, driven by mutation, or more likely epigenetic/transcriptional, factors. I identified greater fungal burdens in larvae infected with clinical strains compared with those infected with environmental strains; this suggests the increased virulence of clinical isolates is driven by increased growth rate or better survival against phagocytosis *in vivo.*

Differences in the pathogenicity of fungal species by source (clinical versus environmental) has been observed in other fungal pathogens including *Aspergillus fumigatus*[236] .

However, it is also possible that adaptation to infection in the human occurs somewhere in the environment – for example through infection in soil microfauna (free-living amoebae, nematodes etc) or in animals. Cryptococcal infection in animals is well described, although it is not clear how frequently this occurs, and there is not such a strong relationship with a particular species as has been described for *Talaromyces*[20] [237]

Differences in the phenotypes of clinical and environmental isolates have also been observed in the *C. neoformans* VNI population in Zambia and South Africa[173]. Clinical isolates produced larger capsules and grew more rapidly at 37^{o}C than environmental isolates. However, laccase activity (involved in melanin production) was slightly higher in environmental isolates (VNI and VNB molecular types) compared with clinical isolates[219]. A study using Brazilian *C. neoformans* also found differences in *in vitro* thermotolerance between clinical and environmental isolates. Clinical *C. neoformans* grew significantly more rapidly at 42^{o}C than environmental counterparts[238]. I speculate that my clinical isolates survive at elevated temperature better than environmental isolates too because the fungal burden of clinical isolates

was significantly higher than that of environmental isolates in larvae incubated at 37^{o}C.

4.5.3 Dose-dependent pathogenicity

I hypothesized that the inoculum size of *C. neoformans* correlates with mortality of larvae. I found denser inoculums of *C. neoformans* resulted in increased probabilities of larval death (Figure 4-4 & 4-5). This effect was more marked in the clinical isolate compared with the environmental isolate. The clinical isolate BMD761 was around 30-fold as potent as the naive environmental strain LD2 in relation to the dose that was associated with a 50% probability of larval death. My finding in larvae is consistent with that in mice. A dose-dependent effect of *C. neoformans* var. *neoformans* in the mouse model has been observed before using clinical isolates, with median survival time negatively correlating with dose [122]. Inoculum size – dependent mortality has also been observed with *C. neoformans* var. *grubii* and *C. gattii* infections in the mouse model with dose varying from 10^3 to 10^6 CFU [239]. An increasing mortality with increasing dose of yeast seems biologically plausible, and is consistent with my findings of higher fungal burdens in larvae infected with more virulent strains.

4.5.4 Larvae-passaged *C. neoformans*

Previously it has been suggested that passaging isolates through the mouse or even in the laboratory can increase virulence of *C. neoformans* H99 strain [240] [241]. I hypothesized that environmental isolates were less virulent than the clinically derived isolates because they had not experienced the immune environment before, and that they would upregulate their virulence if they had sufficient exposure to *Galleria*. Therefore I repeatedly passaged naive isolates through the *Galleria larvae*. However, I found that the virulence characteristics of my strains, both clinical and environmental, were fixed and unaltered by repeated passage through *Galleria* or on agar. This is consistent with a more recent publication that passage of two environmental *C. neoformans* up to three times in mice in fact has little effect on virulence [62]. While I couldn't demonstrate upregulation of virulence solely as a result of passage through *Galleria*, it may be that the period within the *Galleria* was

insufficient to allow such adaptation to occur. *C. neoformans* is likely to be dormant in humans for years before reactivation[72] [62]. An alternative explanation is that the virulence up-regulation is driven by interactions with the adaptive immune system in more complex hosts than *Galleria*. However, my experiments certainly suggest that at some point there is an inter-yeast signaling pathway that becomes active and is likely responsible for maintaining or further enhancing virulence.

4.5.5 Intercellular signaling regulates virulence of *C. neoformans*

I hypothesized that intercellular signaling enables up-regulation of virulence of environmental isolates. My experiments appear to suggest the presence of an inter-yeast signaling pathway that regulates the virulence of *C. neoformans*. Experiments with heat, proteases and nucleases imply that this is likely mediated by secreted peptides/proteins. Previously, it has been reported that secretory vesicles are produced by *C. neoformans* (serotype A and D) strain, and that they contain a number of proteins, many of which are relevant to virulence (growth at 37^{o}C, protection against oxidative stress, capsule synthesis, cell cycle regulation)[119]. More recently, such vesicles have been described for *C. gattii* where they appear to mediate virulence, although the exact mechanism is unclear[230]. Such vesicles are a plausible explanation for the transmissible upregulation I have described here.

Just as quorum sensing has been described in other micro-organisms, signaling peptides that have auto-inducing properties and regulate density-dependent growth have been described in *C. neoformans*[242]. Quorum sensing is the phenomenon whereby microbes coordinate their behaviour by adjusting gene expression in response to population density. In bacteria, this coordination ensures that metabolically expensive virulence factors/processes are only synthesized in synchrony to maximize bacterial survival /advantages[243]. In fungi, the first quorum-sensing molecule (QSM) characterised was farnesol, which mediates morphology switching of *Candida albicans*[244]. *C. albicans* is dimorphic. At densities lower than 10^6 cells/mL, the fungi develop as mycelium; they switch into yeast forms when the density >10^6 cells/mL. Interestingly, farnesol has detrimental effects on other species. It inhibits growth of *Saccharomyces cerevisiae* and induces apoptosis in

Aspergillus nidulans and A. fumigatus[242]. Quorum sensing has also been described in *Histoplasma capsulatum* and *Neurospora crassa*[242].

Quorum sensing (QS)-like molecules were first detected in *TUP1* mutants of *C. neoformans* serotype D when it was observed that inocula of less than 10^3 cells failed to form colonies on agar. This density-dependent growth phenotype could be rescued when *TUP1* mutants were grow in media supplemented with high cell density *TUP1* culture filtrate.[245] This lead to the discovery of an 11-amino acid peptide Qsp1 (quorum sensing-like peptide 1) acting as a signaling molecule in high-density *TUP1* cultures. However, this effect was not observed in *TUP1* strains (serotype A). Ten years after the discovery of Qsp1, the molecular mechanism of this peptide in *C. neoformans* was identified. The *QSP1* gene encodes a peptide precursor consisting of a signal sequence and 24-amino acid pro-peptide (proQsp1). proQsp1 is synthesized intracellularly. Once exported outside, it matures through cleavage mediated by the cell-associated protease Pqp1. The mature Qsp1 enters surrounding cells via peptide transporter Opt1, where it initiates morphological and virulence changes. The virulence of *QSP1*-deficient mutants was attenuated in the mouse inhalation model compared to wild-type H99. The *QSP1* mutants formed wrinkled colonies while H99 strains formed smooth colonies. Moreover the mutant strains were hypomelanized in saturated culture. These defects can be rescued by adding synthetic Qsp1 or placing mutants near WT cells. In saturated culture, transcriptome profiles were significantly different between WT and *QSP1* mutants. Qsp1-regulated genes were enriched for genes associated with cell wall biosynthesis [246]. However, it is not clear that the observation I see here in my experiments is quorum sensing in the sense of being driven by the yeast population density. The exact stimulant that first initiates upregulation of pathogenicity is beyond the scope of this thesis but would be an exciting area for future study.

As mentioned above, a recent report related to the Vancouver Island outbreak strain *C. gattii* (R265) revealed that the strain can accelerate intracellular proliferation rate of non-outbreak strains in murine macrophages. This is regulated by extracellular vesicles (EV) derived from the outbreak strain. EVs are taken up by the host macrophage, then transported to the fungal phagosome where they trigger

proliferation of non-outbreak cells. The two EV components essential for inducing virulence of non-outbreak strains are protein and RNA[230]. However, in my experiments, only protease treatment, rather than nuclease treatment, resulted in loss of the inducing property of culture filtrate.

I examined four ST5 strains derived from HIV-uninfected patients. All of them exhibited induction effects, in contrast to ST5 strains derived from HIV infected patients, where the change in virulence did not meet conventional levels of statistical significance, although the HR was in the direction of an increase. Given that ST5 strains from HIV uninfected patients are distributed throughout the lineage in the phylogenetic tree, I would expect all ST5 isolates to have the potential to infect immunocompetent patients and to be able to be induced, or induce, further strains. Induction effect of some ST5 strains can be explained by phenotypic heterogeneity of ST5 lineage. The ST5 lineage displayed variation of virulence phenotypes *in vitro* compared to non-ST5 strains[247]. Phenotypic heterogeneity facilitates adaptation to environmental fluctuations[248]. I found only the BMD761-induced LD2 strain was able to maintain significantly induced virulence over generations. Moreover, the exclusively hypervirulent BMD761 was able to induce other hypovirulent clinical isolates, including of other sequence types, suggesting virulence is not necessarily associated with sequence type but rather related to distinct attributes of individual strains.

BMD761 in larvae displayed wider variation in fungal burden than LD2 and induced LD2 (Figure 4-19). The burden of induced LD2 in *Galleria* at death was significantly higher than naive LD2. Possible explanations include that induced LD2 grew faster, more efficiently, or evaded phagocytosis better than naive LD2.

Explanations other than peptide/protein signaling could drive the phenomenon of induction. These include viral infection – mycoviruses are described amongst fungal plant pathogens[249], or even prions. Fungal prions are self-propagating and transmissible proteins isoforms, acting as epigenetic determinants by changing their conformations. Prions are proposed to create heritable phenotypic diversity to enable yeast adaptation to fluctuating environments[250 251]. However, the fact that the induction effect was inhibited by boiling and protease treatments suggests that

these latter explanations are unlikely. Hence, the effect of heat and protease treatment suggests that one or more proteins are implicated. Regulation of protein expression can be investigated through transcriptional profiling of protein-coding genes. Regulated genes involved in upregulation of virulence can be validated using real-time reverse transcription PCR. Then potential virulence genes could be precisely knocked-out using CRISPR-Cas9 and validated in animal models. Secretome profiling of the culture filtrate (naive strains, induced strains and knockout strains) could be achieved using liquid chromatography-tandem mass spectrometry.

The changes in virulence phenotype seen in induced isolates cannot be due to recombination. First, I used culture filtrate from the donor inducing strain. Yeasts are too large to pass through the filter. I confirmed the sterility of all culture filtrates for every experiment. Second, I used only purified clonal isolates for every experiment, and finally all culture was done in YPD medium which does not induce mating. Mating in *Cryptococcus* is a rare event that requires specific media[252]. Reversibility of induction can be tested by growing virulent isolates into culture filtrate of naive isolates to see if their virulence phenotypes are suppressed.

Other innate immunity models could be used to test induction of virulence is amoeba, nematode, macrophage and neutrophils. The macrophage model has been widely used as an *in vitro* cellular model of cryptococcal infection[160]. For example, the intracellular proliferation rate in murine macrophages of the non-outbreak *C. gattii* isolate was enhanced by co-culture with the outbreak *C. gattii* isolate[230]. Other animal models that could be utilised include the rat and mouse. Both these species have innate and adaptive immunity, and thus in theory are more similar to humans. Infection is achieved via nasal inhalation or intravenous or intraperitoneal injection. My experimental approach would be essentially the same, first demonstrating that the virulence phenotype is fixed over repeated passage, followed by comparative survival experiments with induced and naïve strains as performed in the *Galleria* model. Rats were reported to develop chronic pulmonary cryptococcosis in the wild, similar to disease due to *C. gattii* in immunocompetnt humans, so it is potentially a realistic model[160]. However, a significant drawback is the relatively large numbers (N=20 per arm) of animals that would be needed.

4.5.6 Morphology of *C. neoformans* post-inoculation in larvae

Surprisingly, I observed pseudohyphae in the induced LD2 population but not in the highly virulent BMD761 isolate. Moreover, pseudohyphae cells were not frequent. They were not always observed in induced LD2 strains. Pseudohyphae have been reported to enhance cryptococcal survival in amoebae, a presumed natural predator of *Cryptococcus* found in soil. The pseudohyphae enable survival by resisting phagocytosis [124] [23]. However, the clinical significance of pseudophyphae is not clear. They failed to kill mice when injected intracranially[253]. No viable cells were revived 28 days post-inoculation. This suggests pseudohyphae are unable to survive within the CNS [253]. Pseudohyphae may not mediate the increased virulence of induced LD2 in *G. mellonella* because they were not always observed in induced LD2 strains, nor did I see them in BMD761 infections.

C. neoformans manifests morphological changes including capsular enlargement and formation of giant cells during mouse pulmonary infection [254] [127]. Cell wall thickness and capsule size increase as a function of infection time in this model. I found that a heterogeneous population (in terms of encapsulated yeast cell, titan cells and pseudohyphae) also develops during infection of Galleria, supporting the validity of using this model to study morphogenesis of *C. neoformans* [167]. I found that induced LD2 isolates had significantly less capsule than naive LD2. However, while large capsule size inhibits phagocytosis by hemocytes, it does not predict survival - survival of larvae infected with large-capsule cells was not significantly different from those infected with small-capsule cells[167]. In my case, the faster growth rate of induced isolates potentially played a vital role in enhanced virulence compared to naive LD2. Rather, the mortality rate of larvae is correlated with fungal load [98]- the higher fungal burden, the higher the mortality rate.

Unexpectedly, I did not find differences in classical virulence traits (thermotolerance, melanin production, CSF *ex vivo* survival) between naive LD2 and Induced LD2, suggestimg that these are virulence rates necessary across all lineages of *C. neoformans*. However, the (probable) faster growth rate of induced LD2, manifest as increased fungal burden, may play a central role in increased virulence, compensating for expression of classical virulence factors.

4.5.7 Comparative pathogenicity of *C. neoformans* and *C. gattii*

Using the *G. mellonella* infection model, I found *C. neoformans* strains derived from immunocompetent individuals were more pathogenic than *C. gattii*. This somewhat surprising result agrees with the finding by Capilla *et al*, and perhaps represents the relative importance of the innate immune system for each of these two infections. Capilla *et al* compared the pathogenicity of both *Cryptococcus* species in both immunosuppressed and immunocompetent BALB/c mice, finding that *C. neoformans* are more virulent than *C. gattii* in both immune states of that mouse[239]. The same result was also seen in the mouse model of severe combined immunodeficiency (SCID).

4.6 Conclusions

My data shows that *G. mellonella* is a good model to compare the pathogenicity of *C. neoformans* strains derived from both HIV-uninfected and HIV patients. My finding supports the hypothesis that ST5 strains, at least those derived from HIV uninfected patients, have increased virulence compared with non-ST5 strains. I also found that environmental isolates are significantly less virulent than clinical isolates.

I found that the virulence of *C. neoformans*, at least for ST5 strains, seems to be regulated by a protein or peptide based signaling system, but the exact nature of this remains unclear. Peptides/proteins may mediate this signaling pathway to enhance virulence of naive environmental isolates, and to coordinate yeast growth during infections. While having a dramatic effect on ST5 strains with lower virulence phenotypes, exemplified by environmental strains, the hypervirulent BMD761 strains also seemed to be able to upregulate the virulence phenotype of strains of different lineages – namely ST4. The enhanced pathogenicity of clinical isolates or induced isolates compared to environmental isolates is associated with higher fungal load in larvae, rather than increase in capsule size or other previously recognised virulence factors. Further experiments with ST5 induced ST4 strains are warranted to identify whether such strains could also upregulate virulence. However, since ST4 strains are rarely found in apparently immunocompetent patients, this is unlikely.

Chapter 5

Comparative transcriptome analysis of environmental and clinical *C. neoformans* var. *grubii* in Vietnam

5.1 Introduction

In chapter 4 I found that:

1. Clinical *C. neoformans* strains are significantly more virulent than MLST-matched environmental isolates in the *G. mellonella* infection model.
2. ST5 clinical strains are more virulent than ST4 clinical strains
3. The virulence state of an isolate is relatively fixed and not altered by repeated culture in *Galleria* or on agar
4. Culture filtrate from ST5 strains from HIV uninfected patients can upregulate the virulence of environmental ST5 strains
5. Culture filtrate from such upregulated ('induced') environmental strains can itself upregulate naïve environmental strains
6. The upregulated virulence state is stable over generations.
7. The inducing property of culture filtrate is inhibited by heat and protease, but not nucleases
8. Culture filtrate from naive environmental strains themselves (i.e never induced) does not upregulate the virulence of naive environmental strains.

This phenomenon of induction of virulence phenotype in *Galleria* may explain the ability of the ST5 lineage to cause disease in the apparently immunocompetent host. There are a number of possible explanations for how the phenomenon of induction is mediated. My experiments in the previous chapter suggest that the effect is mediated by one or more soluble factors; these are unlikely to be nucleases since the effect of induction was not affected by nuclease-treated culture filtrate. Involvement of prion was excluded because the culture filtrate lost effect when exposed to heat or protease. However, the effect of heat and protease treatment suggests that one or more proteins are implicated. In this part of the thesis I use RNA-Seq to answer the following questions:

1. What are the differences in gene expression between environmental and clinical isolates?
2. Is induction associated with upregulation of recognized virulence determinants?
3. What are processes/pathways associated with enhanced virulence of clinical and induced strains?
4. What are gene candidates that may drive the induction?
5. What are their likely functions?

Investigating virulence factors, potential therapeutic targets, of *C. neoformans* is of great interest. Steen *et al* were interested in transcriptional response to survival and progression during rabbit cryptococcal meningitis.[255] They employed the serial analysis of gene expression (SAGE) to characterise transcriptomic profiles of *C. neoformans* isolated from rabbit CSF. SAGE involves constructing and sequencing a library of 10-14 bp tags, whose sequences correspond to unique transcripts. Therefore, the abundance of tags corresponds to the expression level of genes in the cells under specified experimental conditions. Steen *et al* identified several functional pathways important for pathogenicity in rabbit hosts compared to *in vitro* conditions by quantifying tags assigned to genes encoding proteins responsible for protein synthesis and catabolism, stress response, cellular respiration, signal transduction, and transportation as well as carbohydrate, amino acid and lipid metabolism. Cryptococcal cells require these functions to adapt to the rabbit central nervous system.

Hybridization-based microarrays have been used to investigate the transcriptome of *C. neoformans*. Fan *et al* were interested in the transcriptional response of *C. neoformans* to phagocytosis in murine macrophages [95]. They identified that multiple transmembrane transporter genes associated with hexose, amino acids, lipids and iron were induced upon phagocytosis. This increased production of transporters may facilitate the uptake of nutrients by the yeast in the relatively nutrient-depleted phagosome. Macrophage internalization also triggers upregulation of genes related to oxidative and nitrosative stresses. Genes encoding oxidoreductase, peroxidase and denitrosylase were found to be enriched. Interestingly, autophagy genes were

upregulated. Autophagy drives degradation of cytoplasmic components to supply ingredients for recycling of cellular components in the case of nutrient starvation. Furthermore, the association between virulence and mating type was confirmed with the finding that 50 % of MATα locus genes present in the microarray were up-regulated at least twofold. However, further genes related to translation were repressed, such as those encoding ribosomal proteins, translation initiation factors, translation elongation factors and tRNA synthetases. The repression of translation apparatus has also been described in *Candida albicans* following phagocytosis [256]. The response likely reflects a reaction to the nutrient-poor phagosome.

The ability of *C. neoformans* to survive mammalian macrophages has likely been selected through their interaction with soil predators including amoeba, slime molds, and nematodes. This postulation is supported by Derengowski *et al's* study [33] using DNA microarray. They compared transcriptome profiles of *C. neoformans* cells after ingestion by the amoeba *A. castellani*, and murine macrophages. 111 common genes were similarly modulated in both intracellular environments, suggesting their conserved adaptation to evading phagocytosis.

Hybridization-based and sequence-based approaches have been widely used to elucidate transcriptome and explore virulence attributes of *C. neoformans*. However, they have several limitations. Hybridization-based methods require *a priori* knowledge of sequences of interest, are prone to cross-hybridization artefacts and have limited ability to detect highly induced or lowly expressed genes due to saturation signals [257]. Unlike microarrays, tag-based methods can quantify expression level more precisely. However such methods are constrained by the laborious bacterial cloning of sequence tags, the high cost of Sanger sequencing and the requirement for large amounts of RNA input. Furthermore, DNA microarray and tag-based methods cannot detect alternative splicing events or novel transcripts [257].

RNA-Seq (RNA sequencing) is a high-throughput next-generation sequencing (NGS) technology that has revolutionized the field of transcriptomics. RNA-Seq consists of serial experiments including extracting mRNA, reverse transcription to from cDNA, ligation of adapters and construction of a sequence library prior to entry to the NGS platform (Figure 1). It has been widely used to profile the transcriptomes of a variety

of organisms: *S. cerevisiae, Pseudomonas spp*, dog, mouse and human cells [258–260]. In addition to its high sensitivity and accurate measurement of gene expression, RNA-Seq has the ability to detect novel transcripts, alternative splice sites and allele-specific expression.[261]

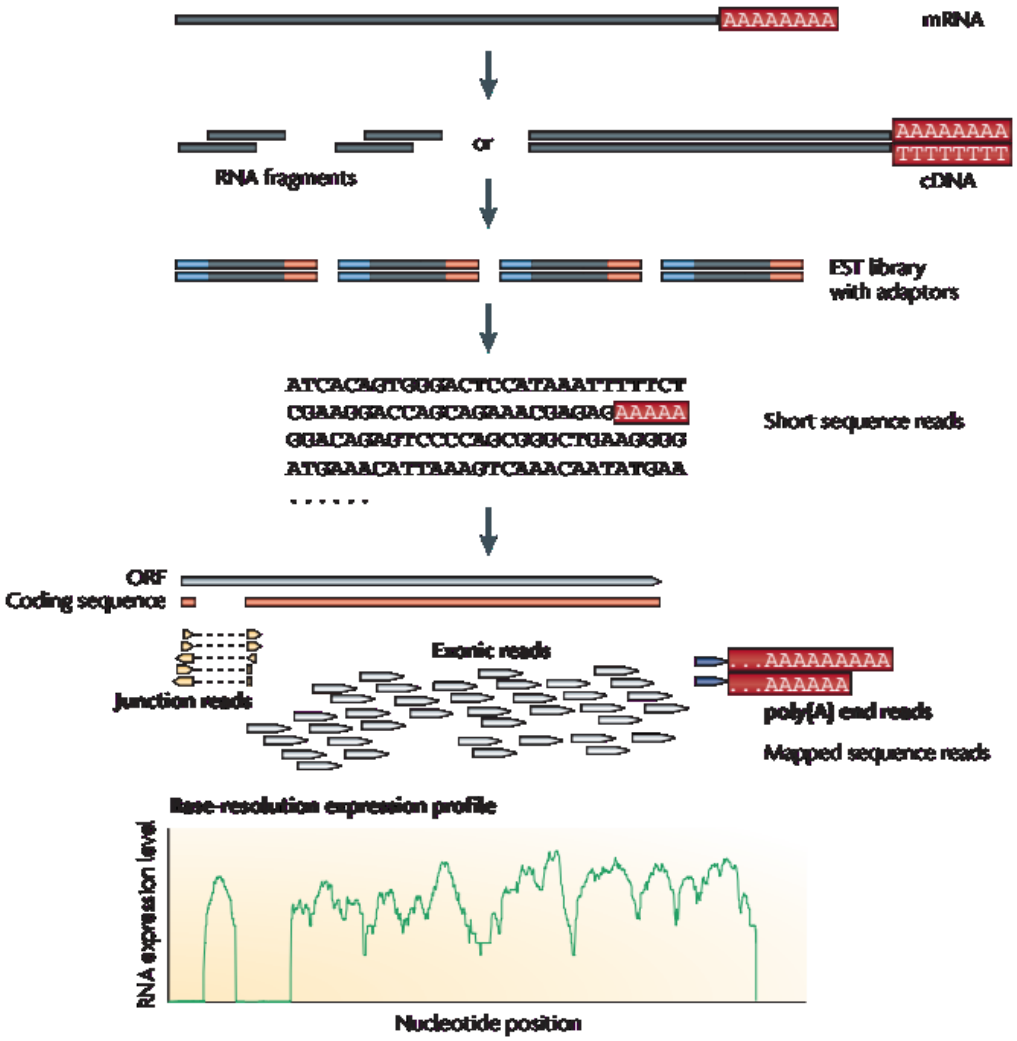


Figure 5-1 Overview of RNA-Seq. First mRNA is fragmented and converted to cDNA by reverse transcription. Second, sequencing adapters (blue) are ligated to each cDNA fragment ends. Next, they are subjected to NGS platform to produce short reads (single end or pair pair end), ranging from 30-400 bp. These reads, including junction reads, exonic reads and polyA-containing reads, are aligned to reference genome or transcriptome. Their alignment to genomic features is counted.

Recently RNA-Seq was used to compare transcriptional profiles of *C. neoformans* var. *grubii* of molecular types VNI and VNII isolated from the CSF of HIV patients with cryptococcal meningitis[262]. These strains were exposed to three environmental

conditions: *in vivo* CSF, *ex vivo* CSF and YPD broth. Surprisingly, the transcriptome profiles of strains grown in *in vivo* CSF and nutritionally replete medium (YPD) were more similar to each other than to those from strains cultured in *ex vivo* CSF. The authors identified 6 genes uniquely upregulated in *in vivo* CSF in both strains. Of 20 genes significantly induced in CSF conditions compared to YPD, some had been previously identified as virulence factors, such as *CFO1, ENA1, RIM101*.

Janbon *et al* employed RNA-Seq to improve annotation of the H99 reference genome [259]. They found that the *C. neoformans* genome is intron-rich [263] - almost all expressed genes harboured at least one with an average of six introns per gene, resulting in a total of 42,000 introns. Introns are important regulators of gene expression in *C. neoformans*, with the accumulation of mRNA of some genes (for example, *CAS3*, *CAP10*, *UGE1, CAS4*) being highly intron dependent[264]. The more introns there are in *CAS3*, the greater the gene induction. Alternative splicing events were also common, identified in 741 genes (more than 10%).

My identification of a difference in virulence between clinical and environmental isolates of *C. neoformans* is not unique – a similar phenomenon has been described before in the mouse model for both *C. neoformans* and *C. gattii* [62 52]. However, these studies did not identify the induction of upregulation of virulence that I have shown in the previous chapter. The induction phenomenon provides me with an elegant biological model where I have congenic strains in the before and after induction state. This offers a unique opportunity to define the transcriptional drivers and markers of virulence using RNAseq.

5.2 Aims

The aims of this chapter are:

- To define the transcriptional differences associated with variability in virulence
- To identify gene candidates associated with variability in virulence between isolates.

5.3 Methods and Material

5.3.1 Strains and culture

The strains used in this chapter are the same strains for the experiment of virulence induction (section 4.4.5.2). I used BMD761, a strain from an HIV uninfected apparently immunocompetent patient of MLST ST5. This patient had been enrolled in a prospective descriptive study of patients with CNS infections [180]. The environmental strain used was LD2, isolated from the bark of a *Hopea odorata* tree from central Ho Chi Minh City during my environmental sampling study. In this chapter I refer to the uninduced environmental strains as naive LD2. Following upregulation of pathogenicity through growth in culture filtrate from BMD761, I refer to it as induced LD2. Self-induced LD2 is naive LD2 that has been grown in culture filtrate that had naive LD2 cultured in it (Table 5-1).

ID	MLST type	Background	Experiment
BK80	4	Isolate from HIV patient	Comparative transcriptome (clinical vs environmental)
BMD761	5	Isolated from HIV-uninfected patient	Comparative transcriptome: clinical vs environmental, induced vs naive
NT7533	5	Isolated from trees	Comparative transcriptome: clinical vs environmental
LD2	5	Isolated from trees	Comparative transcriptome: clinical vs environmental, induced vs naive
Induced LD2	5	LD2 cultured in BMD761 filtrate	Comparative transcriptome: induced vs naive
Self-induced LD2	5	LD2 cultured in its own culture filtrate	Comparative transcriptome: induced vs naive

Table 5-1 Strains used for comparative transcriptional profiling

Phylogenetic relationship of clinical and environmental *C. neoformans* can be found in the appendix B.

5.3.2 Growth curves of strains used

Three replicates of each strain were cultured in YPD broth for 96 hours at 30^{o}C to determine log phase for RNA extraction. An inoculum of approximately 10^{6} cells was grown in 7.5 mL of YPD broth by agitating (200 rpm). At interval time point post-inoculation (0, 6, 20, 24, 30, 44, 54, 68, 72 and 96 hours), an aliquot of 50 uL was taken for quantification by plating in Sabouraud agar plates. These plates were

exposed to incubation at 30^{o}C for 48 hours. Growth curves were plotted using ggplot package and R version 3.4.0 software (R Foundation for Statistical Computing, Vienna, Austria). Growth rates were determined using the formula growth rate=($log10N_t$-$log10N_o$)*2.303/t-t_o, where N indicates cell concentration at a particular time point; and t indicates the time point). The growth curves are presented in appendix C.

5.3.3 Revival, culture and RNA extraction

Single colonies of each isolate were revived from beads (Microbank, UK) stored at -80^{o}C. Revived isolates (approximately 10^6 cells, counted using a Cellometer) were grown in 7.5 ml of YPD broth by agitating for 18 hours at 30^{o}C. My previous kinetic growth experiments indicated *C. neoformans* isolates of interest are in mid-log phase at this time post-inoculation. RNA was harvested using the RiboPure™-Yeast RNA Isolation Kit (Ambion, USA) according to the manufacturer's instructions. Extracted RNA samples were subjected to quality control using an Agilent 2100 Bioanalyzer (Agilent, USA). The integrity of total RNA was estimated using the RIN (RNA Integrity Number), which ranges from 1 to 10, with 1 being the most degraded. Samples with RIN greater than 7 qualified for RNA-Seq. I prepared 6 biological replicates of every experimental condition; all were sent for sequencing [265]. RNA was eluted in isopropanol and shipped at room temperature to Macrogen (Seoul, Korea) for library preparation and sequencing.

5.3.4 Library preparation and Illumina sequencing

In Korea Paired-end RNA-Seq libraries were constructed using the TruSeq stranded mRNA preparation kit (Illumina, USA). Next, cDNA were ligated with sequencing adaptors and were sequenced on an Illumina HiSeq 2000 platform, generating about 11.2 million reads of 150 bp per sample, equivalent to an estimated coverage of 177-fold. Raw data files were downloaded to the OUCRU server.

5.3.5 RNA-Seq analysis

Raw reads were checked for quality using FastQC (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Adapters were

removed using scythe (https://github.com/vsbuffalo/scythe). The resulting reads were further checked based on Phred score and read length to trim low quality regions using Trimmomatic (http://www.usadellab.org/cms/?page=trimmomatic). Sequenced reads were aligned with the *C. neoformans* var. *grubii* H99 reference genome using HISAT2 (https://ccb.jhu.edu/software/hisat2/index.shtml). SAM (Sequences Alignment Map) files, output of read alignment, was converted to BAM (Binary Alignment Map) files using SAMtools (http://www.htslib.org). Finally featureCounts was used to count reads assigned to genomic features (genes, exons, chromosome locations) on the reference genome[266]. The output table of raw counts per gene for each sample was imported for gene expression analysis using DESeq2 [267]. Library size differences were normalized internally[268]. The overview of the RNA-seq workflow is featured below in Figure 5-2

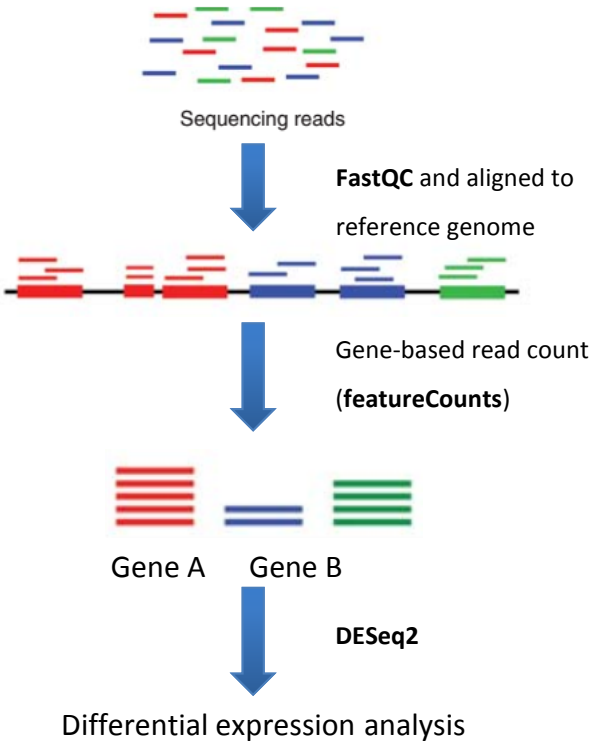


Figure 5-2 Overview of bioinformatic workflow of RNA-Seq data analysis. Following sequencing, raw short reads was subjected to quality control. Next, they were aligned to H99 reference genome. Then abundance of reads assigned to genomic features was determined. Downstreams analyses include differential expression and gene ontology enrichment. Adapted from Kukurba *et al*[257]

5.3.6 Gene ontology (GO) enrichment analysis

Differential gene expression analyses frequently generate datasets with large numbers of genes up- or down-regulated in response to a particular experimental condition. These sets of genes may coordinate to cause system-level changes in the experimental model. To facilitate meaningful biological interpretation for sets of genes, the Gene Ontology Consortium constructed a set of hierarchical controlled vocabulary, allowing the systematic linkage of genes with functions across species [269]. The Gene Ontology vocabularies are constantly curated and updated with over 87,000 species having GO annotations[270]. GO terms have three domains:

- **Biological process:** represents biological objectives to which genes or their products contribute, such as DNA recombination, immune response, signal transduction.
- **Molecular function:** defines biochemical activity of gene products. A product may perform many functions. A set of functions builds up a biological process. Examples of function terms are: catalytic activity, transporter, ligand, DNA binding, signaling.
- **Cellular component:** describes where a gene or a gene product acts within the cell. Examples of this term are nuclear membrane and Golgi apparatus. Multiple gene products can be found in a single cellular component, such as the nucleus or ribosome. This domain reflects understanding of eukaryotic cellular structure, so not every term is applicable to all organisms.

Gene ontology is an important asset in interpreting RNA-Seq data. To this end, I used the R package topGO (available at https://bioconductor.org/packages/release/bioc/html/topGO.html) to test if enrichment of GO terms associated with a set of genes was significantly different from that which would be expected by chance. TopGO uses Fisher's exact test and the Kolmogorov-Smirnov test to test whether GO terms annotated to a set of genes are statistically overrepresented. TopGO takes the GO hierarchy into account to minimize false positive results[271]. The resulting list of significant GO terms was visualised by REVIGO (available at http://revigo.irb.hr). REVIGO reduces functional redundancy and the clustering of semantically similar GO terms[272].

5.3.7 Identification of virulence-associated genes and prediction of hypothetical protein

A list of virulence-associated genes was retrieved from the web page http://www.phidias.us/victors/search_process.php?c_mc_pathogen_id=211 developed by University of Michigan Medical School (USA). Another source for *C. neoformans* virulence genes and prediction of hypothetical protein is the fungal functional genomic database FungiDB (http://www.fungidb.org/fungidb/). Differentially expressed genes are represented by log2 Fold Change which is calculated by averaging for all replicates (treated vs untreated condition).

5.3.8 Experimental workflow

Experimental workflow in this chapter is depicted below

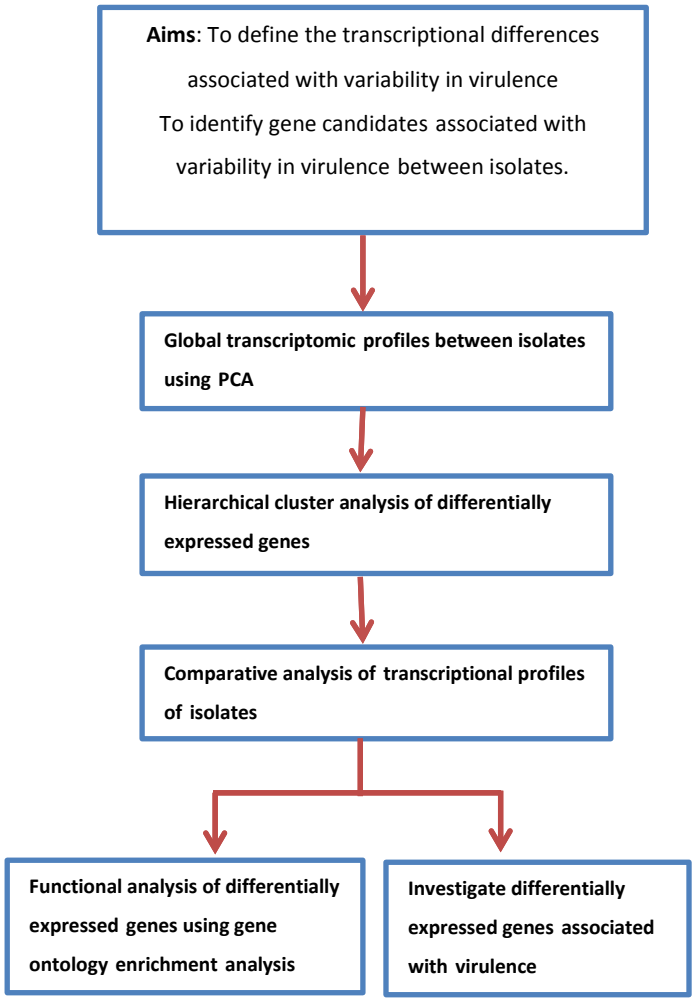


Figure 5-3 Experimental workflow of chapter 5

5.4 Result

5.4.1 RNA-Seq sequencing result

The Illumina HiSeq 2000 platform generated 6.4 to 12.2 million paired-end reads with the length of 150 bp per sample, totaling 331 million reads (Table 5-2). The sequencing depth ranges from 100 to 194X. More than 97% of total reads were mapped to the *C. neoformans* var. *grubii* H99 reference genome, of which at least 94 % were uniquely mapped and taken into account for digital transcriptome analysis.

Experimental condition	Total reads	Mapping (%)	Unique mapping (%)
BK80-1	9694924	98.21	95.76
BK80-2	9345272	99.31	96.75
BK80-3	7828385	99.22	96.67
BK80-4	7906881	99.28	95.33
BK80-5	8895059	99.31	97.03
BK80-6	8288584	99.35	95.11
BMD716-1	10566932	99.07	96.56
BMD716-2	10209953	99.22	94.11
BMD716-3	8026923	99.16	96.47
BMD716-4	8389000	99.02	96.63
BMD716-5	7992309	99.07	96.82
BMD716-6	10283111	99.08	96.75
NT7533-1	6411961	98.68	96.37
NT7533-2	10948992	99.15	96.49
NT7533-3	9427071	99.00	96.53
NT7533-4	11626846	99.04	96.33
NT7533-5	6871780	99.09	96.77
NT7533-6	6467683	99.07	96.73
LD2-1	9483009	99.01	96.63
LD2-2	11023417	99.13	96.81
LD2-3	10871841	98.94	96.49
LD2-4	9038152	97.23	94.84
LD2-5	11368845	99.07	96.60
LD2-6	6879322	99.11	96.90
Induced LD2-1	10297952	99.17	96.96
Induced LD2-2	8650109	99.19	96.76
Induced LD2-3	9502629	99.19	96.33
Induced LD2-4	10857524	99.06	96.67
Induced LD2-5	10559163	99.20	97.13
Induced LD2-6	10979132	99.24	95.95
Self-induced LD2-1	8345534	98.94	96.18
Self-induced LD2-2	7389997	99.31	96.73
Self-induced LD2-3	6901251	99.22	96.54
Self-induced LD2-4	8314394	99.27	96.51
Self-induced LD2-5	9311518	99.00	96.10
Self-induced LD2-6	12251864	99.22	96.42

Table 5-2 Mapping statistics for *C. neoformans* var. *grubii* RNA-seq data. Each isolate was performed in six biological replicates. Total reads: number of short read generated; Mapping: percentage of reads mapped to *C. neoformans* var. *grubii* H99 reference genome. Unique mapping: percentage of reads uniquely mapped to genomic features in the reference genome.

5.4.2 Global transcriptomic profiles between clinical and environmental isolates are different

To examine overall similarity and difference in transcriptome between samples, I used principle component analysis (PCA) to analyse clustering of clinical isolates (BK80 and BMD761) vs environmental isolates (NT7533 and LD2). The data points (24 biological isolates) were projected onto a 2D plane such that they spread out in two dimensions to reveal the most variance. PC1 represents the direction that data points separate the most. In this case, PC1 (principle component 1) explains 62% of the variance in transcriptional profiles between samples (Figure 5-4). In the plot, three groups form 3 clusters by source and MLST genotypes: clinical isolates (ST4), clinical isolates (ST5) and environmental isolates (ST5). Strikingly, the majority of environmental isolates are separated from clinical isolates (regardless of genotype), which agrees with the significant difference in survival phenotype seen on infection of *Galleria* (Chapter 4). There appeared to be two outliers in my dataset: two from the environmental isolate LD2 and one from the clinical isolate BMD761. The reason for these outliers is unclear but may reflect technical variations and/or population heterogeneity. Given that these outliers may have represented a technical variation, and because I could not be certain that the phenotype of the strains in the *Galleria* model had changed, I removed these outliers from further analyses.

After removal of the 3 outliers, clustering was still maintained (Figure 5-5). Clinical isolates were grouped by genotype and were separated from a cluster of environmental replicates. PC1 now explained 65 % of the variance in transcriptome profiles between samples.

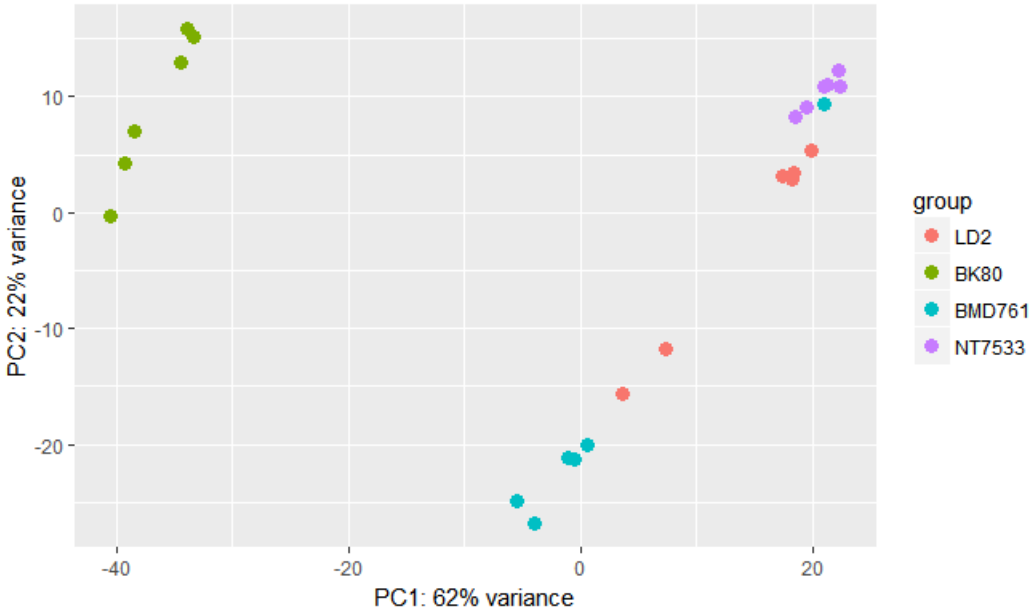


Figure 5-4 Principle component analysis of transcriptional profiles. 4 isolates with 6 biological replicates each are included: BK80, BMD761, NT7533 and LD2. NT7533 and LD2 are environmental isolates. The biological replicates are separated well by genotype (ST4 vs ST5) and source (clinical vs environmental) except three outliers (one BMD761 and two LD2 replicates).

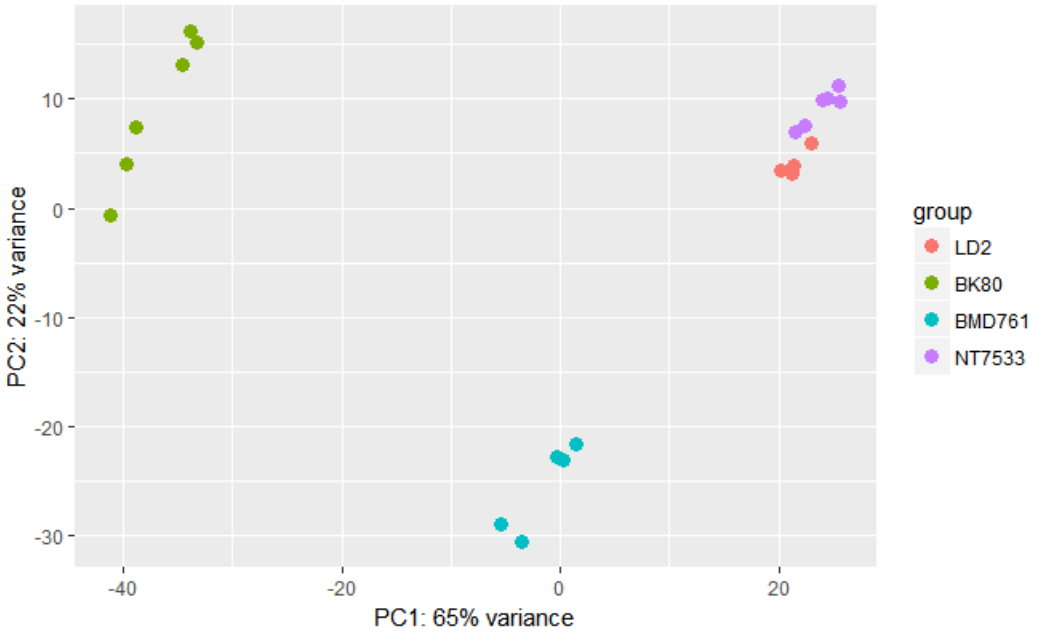


Figure 5-5 Principle component analysis of transcriptional profiles after removal of 3 outliers. Red dots and blue dots indicate environmental isolates and clinical isolates, respectively. The biological replicates are separated well by genotype (ST4 vs ST5) and source (clinical vs environmental).

5.4.3 Comparative analysis of transcriptome between clinical and environmental isolates

Comparative analysis of clinical vs environmental isolates (BK80 plus BMD761 vs LD2 plus NT7533) was performed to test for differentially expressed genes. A total of 1701 genes were differentially expressed between the two isolate groups (Wald test, Benjamini-Hochberg adjusted P value < 0.05, log2 fold change ≥ 1 and ≤ -1). A hierarchical clustering of 1701 differentially expressed genes based on normalised read count per gene reveals two source-specific gene clusters. The gene expression patterns of clinical isolates is different from that of environmental isolates (Figure 5-6) with 977 genes induced and 724 genes repressed in clinical isolates relative to environmental isolates. In addition, differences in gene expression by sequence type were evident in clinical isolates (BK80 (ST4) vs BMD761 (ST5)). NT7533 and LD2 are all sequence type 5, but their transcriptome patterns differ. The heat map also shows some variation (technical and/or biological) in expression pattern within biological isolates (BK80 and BMD761, Figure 5-6).

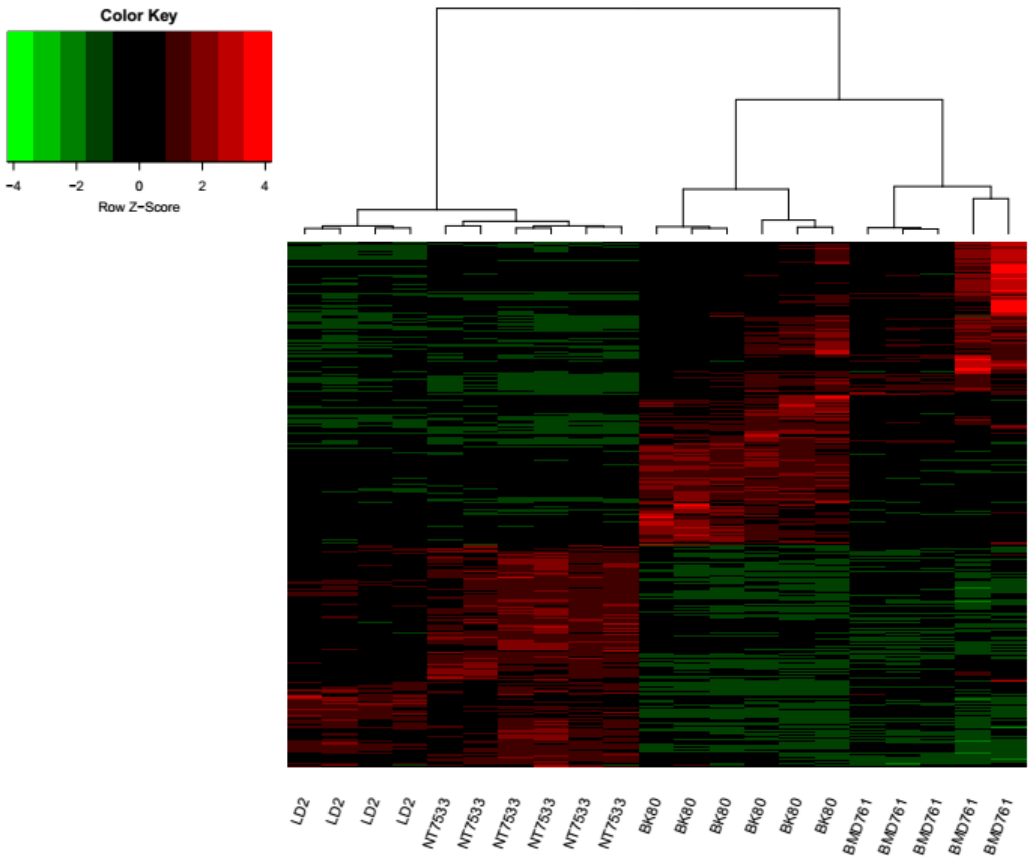


Figure 5-6 Hierarchical cluster analysis of gene expression based on nomalised read counts per gene. The heat map displays expression patterns for 21 samples. Each sample corresponds to 1 column and each single gene to a single row. The colour key ranges from red for up-regulated gene to green for down-regulated genes across samples.

5.4.4 Functional analysis using gene ontology

5.4.4.1 Biological process terms relevant to cellular biosynthesis and metabolism were enriched in clinical isolates vs environmental isolates

The set of up-regulated genes in clinical isolates relative to environmental isolates was investigated for overrepresentation of GO term annotations in three domains (biological process, molecular function and cellular component). A list of significantly enriched GO terms associated with biological process was summarised and visualised using the semantic similarity REVIGO scatterplot (Figure 5-7 & Table 2). In the

bubbles representing GO terms, the colours correspond to the significance level (P value < 0.05) and the size indicates the frequency of GO terms in the underlying Gene Ontology annotation database (UniProt). The smaller the bubble, the more specific the GO terms. The shorter the distance between bubbles, the more the semantic similarity.

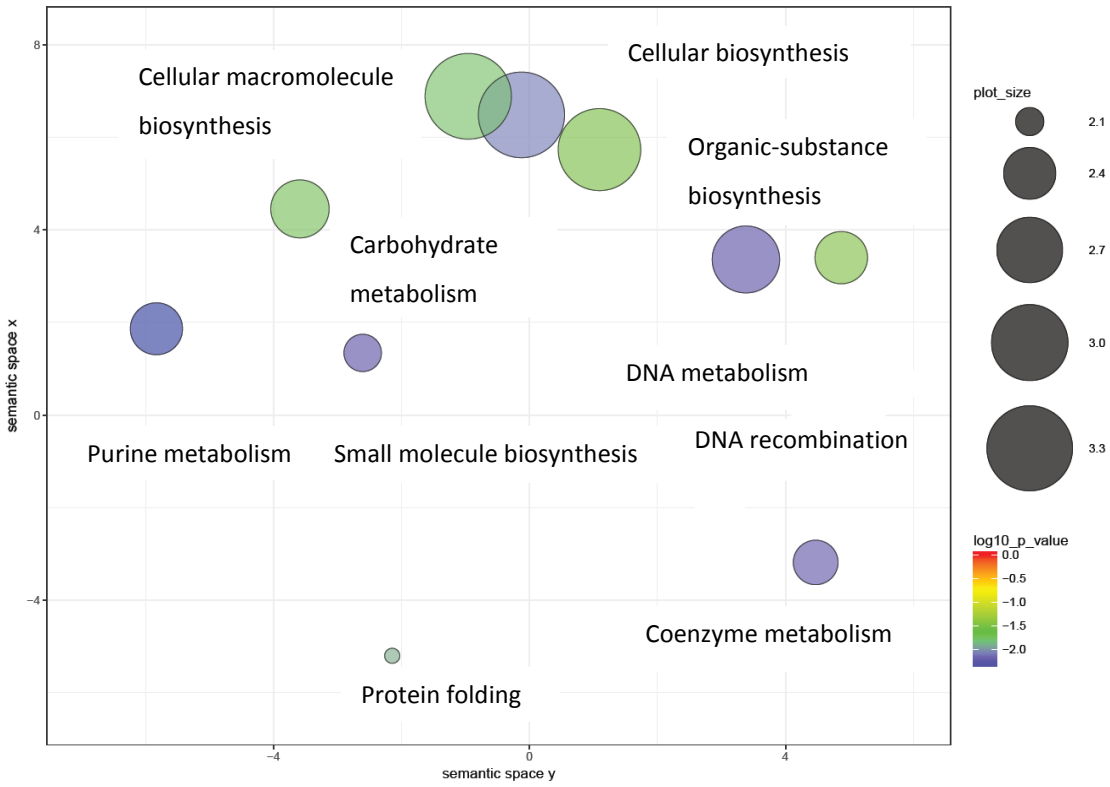


Figure 5-7 Biological process–associated GO terms of upregulated genes in clinical isolates. Bubble size represents frequency of GO terms in the GO annotation database (UniProt). Colour indicates significant level. Bubble proximity implies their semantic similarity.

Strikingly, the most significantly enriched GO terms were related to cell metabolism: purine ribonucleotide metabolism (GO:0009150), coenzyme metabolism (GO:0006732), and DNA metabolism (GO:0006259). Upregulated genes involved in biosynthesis were frequently enriched: cellular macromolecule biosynthesis (GO:0034645), organic substance biosynthesis (GO:1901576) and cellular biosynthesis (GO:0044249) (Table 5-3). A gene participating in protein modification is also upregulated (GO:0006457, protein folding).

GO term ID	Description	Frequency	log10 p-value
GO:0009150	purine ribonucleotide metabolic process	2.26%	-2.1871
GO:0006457	protein folding	1.69%	-1.8996
GO:0006732	coenzyme metabolic process	2.81%	-2.1675
GO:0005975	carbohydrate metabolic process	4.50%	-1.5331
GO:0044249	cellular biosynthetic process	31.29%	-2.0605
GO:0006310	DNA recombination	3.81%	-1.4881
GO:0006259	DNA metabolic process	8.23%	-2.1871
GO:0044283	small molecule biosynthetic process	4.99%	-1.7447
GO:0034645	cellular macromolecule biosynthetic process	23.36%	-1.5361
GO:1901576	organic substance biosynthetic process	31.65%	-1.767

Table 5-3 Enriched biological process–associated GO terms of upregulated genes in clinical isolates. Given are GO term ID, term descriptions, frequency of GO terms in the underlying GO (UniProt) and log10 P value.

5.4.4.2 Molecular function terms relevant to protein biosynthesis and cell division were enriched in clinical isolates

Figure 5-8 displays enriched GO terms associated with molecular function. The most significantly enriched GO term related to protein synthesis machinery: structural constituent of ribosome (GO:0003735) (Table 5-4). Other significantly enriched GO terms were those associated with microtubules (microtubule motor activity and microtubule binding). Moreover, genes involved in the synthesis of DNA and polypeptides were enriched: DNA-directed DNA polymerase activity (GO:0003887) and translation elongation factor activity (GO:0003746).

I also analysed enrichment of GO terms for the cellular component domain. Only one GO term was significantly enriched: GO:0000922 spindle pole.

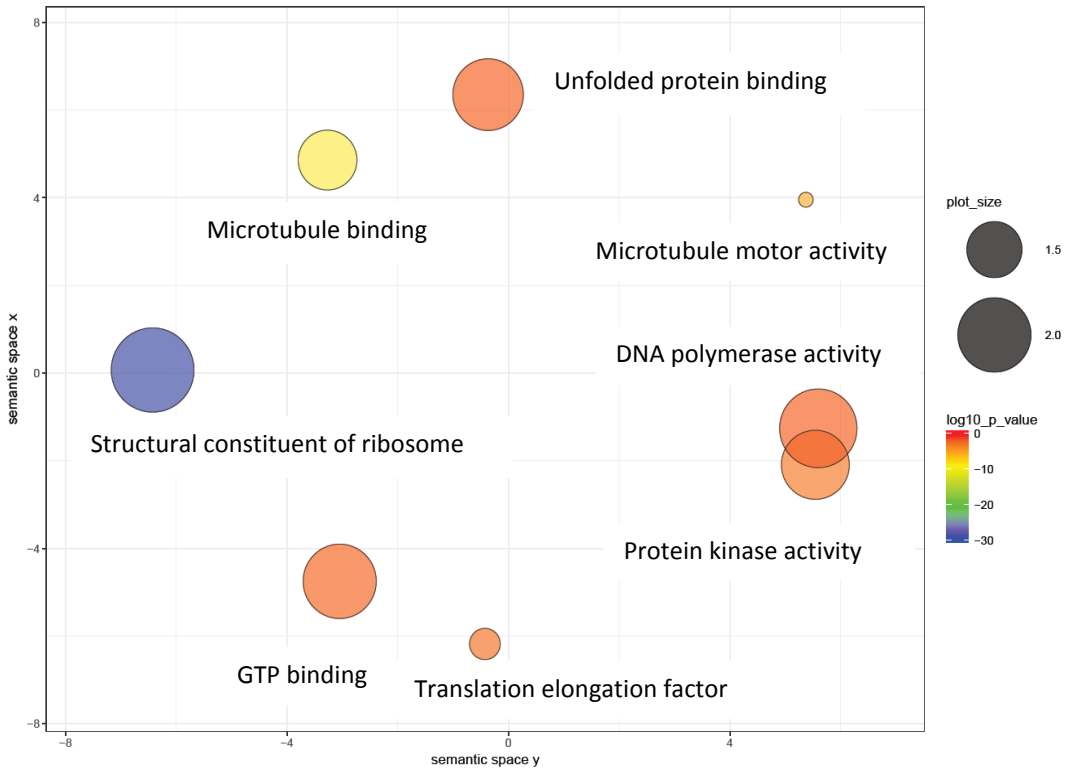


Figure 5-8 Molecular function–associated GO terms of upregulated genes in clinical isolates. Bubble size is proportional to frequency of GO terms in the underlying GO annotation database (UniProt). Colour indicates significant level. Bubble proximity implies their semantic similarity.

GO term ID	Description	Frequency	log10 p-value
GO:0003735	structural constituent of ribosome	3.29%	-30
GO:0003777	microtubule motor activity	0.15%	-5.4685
GO:0008017	microtubule binding	0.57%	-8.6021
GO:0003746	translation elongation factor activity	0.18%	-2.3372
GO:0003887	DNA-directed DNA polymerase activity	1.01%	-2.7696
GO:0051082	unfolded protein binding	1.22%	-1.3851
GO:0005525	GTP binding	1.49%	-1.644
GO:0004672	protein kinase activity	2.07%	-1.3478

Table 5-4 Enriched molecular function–associated GO terms of upregulated genes in clinical isolates.

5.4.4.3 Gene ontology analysis of down-regulated genes in clinical isolates

In contrast to up-regulated genes, there were fewer enriched GO terms due to the low number of down-regulated genes in clinical isolates. For the biological process domain, only two GO terms were significantly enriched: GO:0009152 purine ribonucleotide biosynthesis and GO:0055114 oxidation-reduction process with the latter being more statistically significant than the former (Table 5-5 and Figure 5-9). Notably, half of the molecular-function GO terms were related to transporter activity: transporter activity, nucleobase transmembrane transporter activity, active ion transmembrane transporter activity, and carboxylic acid transmembrane transporter activity (Figure 5-10 & Table 5-5). Additionally, genes involved in catalysis were down-regulated: hydrolase activity and acid phosphatase activity. No genes annotated to the cellular component GO domain were enriched.

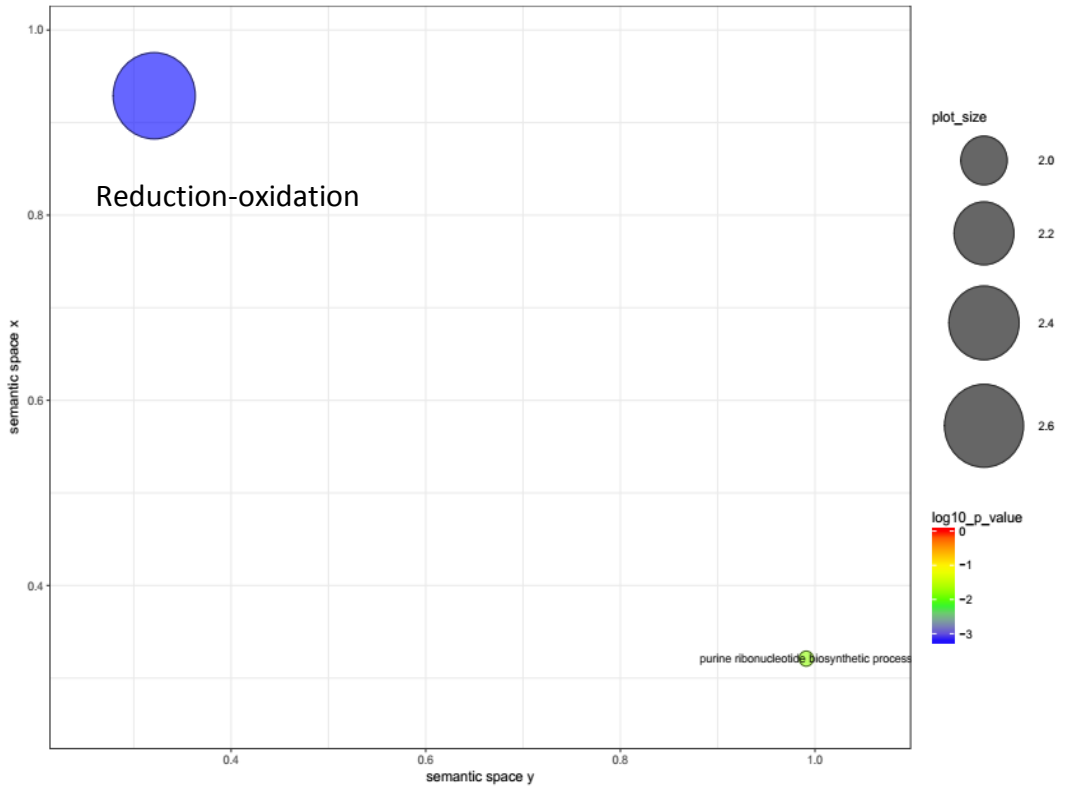


Figure 5-9 Enriched biological process–associated GO terms of downregulated genes in clinical isolates. Bubble size is proportional to frequency of GO terms in the underlying GO annotation database (UniProt). Colour indicates significant level. Bubble proximity implies their semantic similarity.

GO domain	GO term ID	Description	Frequency	log10 p-value
Molecular function	GO:0003993	acid phosphatase activity	0.1%	-1.8
Molecular function	GO:0005215	transporter activity	7.1%	-2.5
Molecular function	GO:0008047	enzyme activator activity	2.1%	-1.8
Molecular function	GO:0015205	nucleobase transmembrane transporter activity	0.1%	-2.4
Molecular function	GO:0016616	oxidoreductase activity, acting on the CH-OH group of donors, NAD or NADP as acceptor	1.2%	-1.3
Molecular function	GO:0016787	hydrolase activity	15.9%	-2.5
Molecular function	GO:0022853	active ion transmembrane transporter activity	0.7%	-1.4
Molecular function	GO:0046943	carboxylic acid transmembrane transporter activity	1.2%	-1.4
Biological process	GO:0009152	purine ribonucleotide biosynthetic process	0.9%	-1.8
Biological process	GO:0055114	oxidation-reduction process	7.1%	-3.2

Table 5-5 Overrepresented GO terms associated with biological process and molecular function in clinical isolates

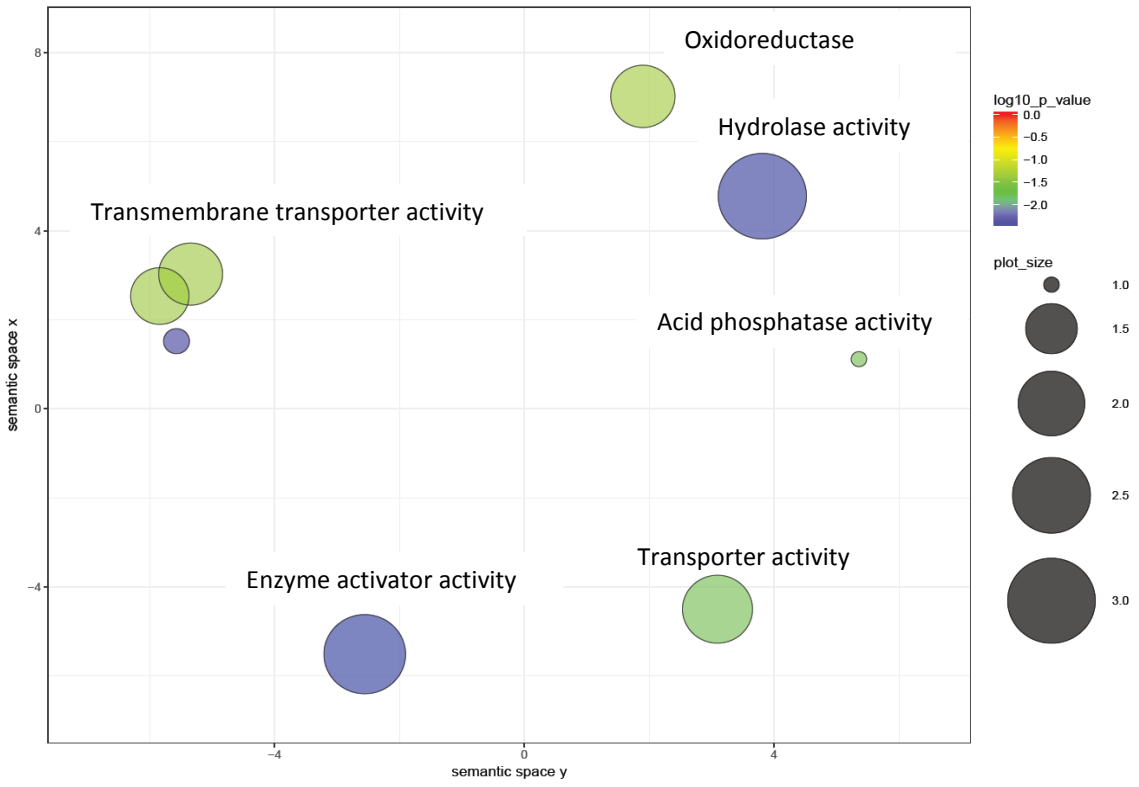


Figure 5-10 Enriched molecular function–associated GO terms of down-regulated genes in clinical isolates. Bubble size represents frequency of GO terms in the GO annotation database (UniProt). Colour indicates significant level. Bubble proximity implies their semantic similarity.

5.4.5 Differentially expressed genes associated with virulence between clinical and environmental isolates

5.4.5.1 Regulation of Gat201 target genes between clinical and environmental isolates

The pleiotropic virulence determinant Gat201, a GATA-family zinc finger DNA transcription factor, is a key regulator of virulence and capsule-independent antiphagocytosis. The virulence of the *GAT201* knockout mutant is attenuated in the murine infection model compared to wild type[226]. Chun *et al* showed that *GAT201* controls the induction of ~1100 genes (16 % of the genome) and can affect a twofold difference in transcript level. Of 1100 genes, 62 genes have promoters bound to *GAT201* for regulation of gene expression in tissue culture conditions[116]. Table 5-6 displays the Gat201-bound genes that are differentially regulated in my clinical compared with environmental *C. neoformans* isolates. Of 19 Gat201 target genes with differential expression, 4 were upregulated and 15 were down regulated at least 2 fold in clinical isolates (Table 5-6).

ID	Product Description	Name	log2FoldChange
CNAG_06186	sugar transport	CNM1740	4.1
CNAG_05867	**L-fucose transporter**	**CNF3650**	**1.4**
CNAG_05664	branched-chain-amino-acid transaminase	BAT1	1.1
CNAG_05312	**macrophage activating glycoprotein**	**CNI3590**	**1.1**
CNAG_05229	stomatin family protein	CNI2770	-1.1
CNAG_05147	**hypothetical protein**	**CNI2090**	**-1.2**
CNAG_04756	ribonuclease H-related protein	CNJ1610	-1.3
CNAG_04737	**potential methyltransferase**	**CNJ1790**	**-1.3**
CNAG_04736	hypothetical protein	CNJ1800	-1.4
CNAG_04735	**metalloprotease**	**MEP1**	**-1.5**
CNAG_04517	hypothetical protein	null	-1.5
CNAG_03122	**hypothetical protein**	**null**	**-1.5**
CNAG_03012	quorum sensing peptide	QSP1	-1.6
CNAG_01601	**potential vacuolar triglyceride lipase**	**ATG15**	**-1.6**
CNAG_01553	CORD and CS domain-containing protein	CNC0890	-1.6
CNAG_00919	**carboxypeptidase D**	**KEX101**	**-1.7**
CNAG_00456	hypothetical protein	null	-1.8
CNAG_00374	**hypothetical protein**	**null**	**-1.9**
CNAG_00165	methylthioadenosine phosphorylase	MEU1	-2.3

Table 5-6 Gat201 target genes differentially expressed in Clinical vs environmental *C. neoformans*

5.4.5.2 Regulation of genes associated with murine lung infectivity between clinical and environmental isolates

A number of genes have been identified in *C. neoformans* that are virulence factors in terms of ability to infect and proliferate in murine lungs[226]. I found some of these genes differentially expressed in clinical versus environmental strains (Table 5-7). Two of them – *PDR802* and *LIV10* - were induced in clinical isolates, and 4 were down-regulated (Table 5-7)

ID	Product Description	Name	log2FoldChange
CNAG_03894	Binuclear cluster DNA-binding regulator	PDR802	1.93
CNAG_02930	**DUF1765 domain protein of unknown function**	**LIV10**	**1.00**
CNAG_04611	ubiquitin-conjugating enzyme	UBC8	-1.18
CNAG_00531	**potassium/sodium efflux P-type ATPase**	**ENA1**	**-1.24**
CNAG_02409	protein of unknown function	LIV5	-1.41
CNAG_00130	**Ca-CaM kinase homolog**	**RCK2**	**-1.45**

Table 5-7 Regulation of genes associated with murine lung infectivity in clinical vs environmental isolates**.**

5.4.5.3 Multiple kinase genes were upregulated in clinical isolates

Kinases play vital roles in regulating multiple processes (growth and cell cycle, nutrient metabolism, stress response and adaptation, cell signaling and many other cellular functions), phosphorylating 30 % of cellular proteins [227].Kinases are important candidates for therapeutic targets as they are possibly inhibited by small molecules or antibodies. A set of kinase encoding genes has been demonstrated pathogenicity in the *G. mellonella* killing assay and murine infection model [227]. Here I detected differential expression of these genes in clinical vs environmental isolates. Seven out of eight of these genes were up-regulated in clinical isolates (Table 5-8).

ID	Product Description	Name	log2FoldChange
CNAG_02675	CAMK/CAMKL/GIN4 protein kinase	HSL101	4.2
CNAG_03184	**BUB protein kinase**	**BUB1**	**4.0**
CNAG_03369	Protein tyrosine kinase	SWE102	3.1
CNAG_02389	**AGC protein kinase**	**YPK101**	**1.8**
CNAG_06697	TTK protein kinase	MPS1	1.6
CNAG_05558	**serine/threonine protein kinase**	**KIN4**	**1.1**
CNAG_05439	CAMK/CAMK1/CAMK1-CMK protein kinase	CMK1	1.1
CNAG_03024	**AGC protein kinase**	**RIM15**	**-1.2**

Table 5-8 Regulation of pathogenicity-related kinase genes in clinical vs environmental isolates

5.4.5.4 Regulation of transcription factors between clinical and environmental isolates

Transcription factors govern general biology and pathogenicity of *C. neoformans* through activating or repressing effectors of multiple signaling pathways, including Ras, calcineurin, MAP kinase, and *RIM101* [29–31]. Such transcription factors have been shown to be enriched when *C. neoformans isolates* grown in pigeon guano were compared with those cultured under starvation or rich medium conditions (YPD). This suggests transcription factors are important in cryptococcal adaptation to different ecological niches[259]. I identified differences in the expression of transcription factor genes between environmental and clinical strains. Of note, the pleiotropic virulent gene *GAT201* (CNAG_01551) was down-regulated in clinical isolates (Table 5-9)

ID	Product Description	Name	log2FoldChange
CNAG_03894	Binuclear cluster DNA-binding regulator	PDR802	1.9
CNAG_01438	**transcription factor**	**MBS2**	**1.8**
CNAG_04457	hypothetical protein	FZC30	1.5
CNAG_05153	**hypothetical protein**	**GAT5**	**1.2**
CNAG_02603	early growth response protein 1	ZFC1	-1.1
CNAG_01551	**hypothetical protein**	**GAT201**	**-1.2**

Table 5-9 Regulation of transcription factors in clinical vs environmental isolates

5.4.5.5 Established virulence-associated genes are upregulated in clinical isolates

C. neoformans can survive and proliferate in animal hosts through interaction between fungal virulence factors and host susceptibility. Virulence factors are all mechanisms that enable the pathogen to survive and evade the host immunity. Cryptococci do not rely on a single virulence factor for pathogenicity but rather require a minimum set of interacting attributes for pathogenicity[32,274,275]. Established virulence factors include genes controlling thermotolerance, capsule enlargement, melanin production, phospholipase activity, oxidative defense, antiphagocytic proteins, nutrient acquisition, morphological switching, etc. I retrieved a list of 186 genes encoding established virulence factors (available at http://www.phidias.us/victors/index.php). From this list, 28 genes were differentially regulated in clinical vs environmental *C. neoformans*; 19 of these genes were upregulated in clinical strains (Table 5-10).

ID	Product Description	Name	log2FoldChange
CNAG_06487	putative chitin synthase	CHS6	5.5
CNAG_01907	**PLK/PLK1 protein kinase**	**null**	**5.2**
CNAG_00779	large subunit ribosomal protein L27e	RPL27	2.2
CNAG_06324	**zinc finger protein**	**null**	**2.1**
CNAG_05274	STE/STE20/YSK protein kinase	null	2.1
CNAG_05968	**rho-like GTPase**	**CDC420**	**2.0**
CNAG_06231	large subunit ribosomal protein L13	null	1.9
CNAG_02531	**mitogen activated protein kinase**	**CPK1**	**1.9**
CNAG_02729	hypothetical protein	FRR1	1.9
CNAG_00418	**S-adenosylmethionine synthase**	**ILV2**	**1.7**
CNAG_00869	ATP-binding cassette transporter	PDR5	1.7
CNAG_01890	**methionine synthase**	**MET6**	**1.5**
CNAG_03326	putative chitin synthase	CHS2	1.4
CNAG_05998	**rho-GTPase**	**RAC2**	**1.3**
CNAG_01172	parallel beta-helix repeat protein	PBX1	1.3
CNAG_02036	**putative sugar transporter**	**CAS4**	**1.2**
CNAG_03263	translation elongation factor Tu	TUF1	1.2
CNAG_00124	**hypothetical protein**	**CAS32**	**1.1**
CNAG_05144	carbonic anhydrase	CAN2	1.0
CNAG_03202	**adenylate cyclase**	**CAC1**	**-1.1**
CNAG_05181	white-collar transcription factor; blue-light photoresponsive gene	BWC1	-1.1
CNAG_01138	**putative mitochondrial cytochrome c peroxidase**	**CCP1**	**-1.2**
CNAG_06050	UDP-glucose 4-epimerase, UDP-glucose 4-epimerase	UGE2	-1.2
CNAG_05449	**hypothetical protein**	**CMT1**	**-1.4**
CNAG_05015	catalase 4	CAT4	-1.5
CNAG_03465	**laccase**	**LAC1**	**-2.2**
CNAG_05472	meiotic recombination protein SPO11	null	-2.3
CNAG_06574	**secreted antiphagocytic protein**	**APP1**	**-2.6**

Table 5-10 Regulation of virulence-associated genes in clinical vs environmental isolates

5.4.5.6 Multiple transporter genes were induced in clinical isolates

Transporters plays important roles in yeast pathobiology[262]. Some are responsible for multiple drug resistance to fungus: major facilitator superfamily (MFS) transporters and ATP-binding cassette (ABC) transporter[276] [277]. Others enable *Cryptococcus* to adapt to different ecological niches, such as pigeon excreta, murine lung and human CSF [259] [278] [201]. Table 5-11 and 5-12 list the up- and down-regulated genes seen in clinical isolates. Twice as many transporter genes were up-regulated as were down regulated in clinical isolates. In particular, multiple genes encoding sugar, zinc ion, amino acid and monosaccharide transporters were exclusively enriched in clinical isolates (Table 5-11). Some MFS and ABC genes were upregulated in clinical isolates (CNAG_00792, CNAG_00869, CNAG_04416, CNAG_01674 & CNAG_05982)

whereas others were upregulated in environmental isolates (CNAG_07799, CNAG_00905 & CNAG_01690) (Table 5-11 & 5-12). Of upregulated transporters in clinical isolates, the monosaccharide transporter (CNAG_02733) was the most highly induced.

GeneID	Product Description	Name	log2FoldChange
CNAG_05833	allantoate transporter	null	1.3
CNAG_06561	**allantoate transporter**	**null**	**1.4**
CNAG_05345	amino acid transporter	null	1.7
CNAG_07367	**amino acid transporter, amino acid transporter**	**null**	**1.6**
CNAG_00792	ATP-binding cassette transporter	null	1.7
CNAG_00869	**ATP-binding cassette transporter**	**PDR5**	**1.7**
CNAG_04567	drug transporter	null	1.9
CNAG_03426	**GDP-mannose transporter 2**	**GMT2**	**1.6**
CNAG_00303	high-affinity nicotinic acid transporter	null	2.2
CNAG_04416	**major facilitator superfamily transporter**	**null**	**1.6**
CNAG_01674	MFS transporter	null	1.7
CNAG_05982	**MFS transporter, SP family, general alpha glucoside:H+ symporter**	**null**	**1.5**
CNAG_02733	monosaccharide transporter	null	3.2
CNAG_05340	**monosaccharide transporter**	**null**	**1.2**
CNAG_00867	myo-inositol transporter	ITR3A	1.4
CNAG_05377	**myo-inositol transporter**	**ITR3**	**1.8**
CNAG_06545	pim1 protein RNA transporter 2	null	1.3
CNAG_02036	**putative sugar transporter**	**CAS4**	**1.2**
CNAG_06761	siderophore-iron transporter Str1	null	1.1
CNAG_07387	**siderophore-iron transporter Str3**	**null**	**1.3**
CNAG_05075	sodium-dependent phosphate transporter	null	3.0
CNAG_01752	**mitochondrial 2-oxodicarboxylate transporter**	**null**	**1.5**
CNAG_03274	mitochondrial S-adenosylmethionine transporter	null	1.4
CNAG_05254	**UDP-galactose transporter**	**null**	**1.2**
CNAG_01936	sugar transporter	null	1.7
CNAG_02586	**sugar transporter**	**null**	**1.4**
CNAG_05324	sugar transporter	null	1.7
CNAG_06253	**sugar transporter**	**null**	**1.7**
CNAG_06932	sugar transporter	null	1.0
CNAG_06963	**sugar transporter**	**null**	**1.9**
CNAG_07730	tricarboxylate transporter	null	1.6
CNAG_00895	**zinc ion transporter**	**ZIP1**	**2.9**
CNAG_03398	zinc ion transporter	ZIP2	2.2

Table 5-11 Up-regulation of transporter genes in clinical isolates

GeneID	Product Description	Name	log2FoldChange
CNAG_07902	AAT family amino acid transporter	null	-1.8
CNAG_07799	**ABC transporter**	**null**	**-2.0**
CNAG_04758	amt family ammonium transporter	AMT2	-2.7
CNAG_02455	**choline transporter**	**null**	**-1.6**
CNAG_00979	copper transporter	CTR4	-1.8
CNAG_05867	**L-fucose transporter**	**null**	**-1.5**
CNAG_06890	membrane transporter	null	-1.9
CNAG_00905	**MFS transporter**	**null**	**-2.6**
CNAG_01690	MFS transporter	null	-2.1
CNAG_04546	**multidrug transporter**	**null**	**-1.6**
CNAG_05994	multidrug transporter	null	-1.2
CNAG_04617	**OPT family small oligopeptide transporter**	**null**	**-1.7**
CNAG_04794	spermine transporter	null	-1.0
CNAG_05993	**transmembrane transporter Liz1p**	**null**	**-1.1**
CNAG_07448	urea transporter	DUR3	-2.7
CNAG_04306	**vesicle transporter SFT2B**	**null**	**-1.0**

Table 5-12 Down-regulation of transporter genes in clinical isolates

5.4.6 Global transcriptomic profiles of clinical BMD761 strain and induced LD2 strain are different from those of self-induced LD2 and naive LD2 strains

I used principle component analysis (PCA) to analyse clustering of clinical isolates (BMD761), environmental isolates (LD2), Induced LD2 and Self-induced LD2 strains. PC1 (principle component 1) explains 64% of variance in transcriptional profiles between samples (Figure 5-11).

The majority of replicates clustered according to their experimental isolates. Some outlier replicates were somewhat intermingled including three BMD761, one induced LD2 and two naive LD2 replicates. The emergence of outliers may reflect technical variations and/or population heterogeneity. Technical variation is inevitable in any complex experiment and in this case, this may have arisen from library preparation or sequencing (lane effect, flow cell effect) or both. Biological variation is intrinsic to all organisms. Even genetically identical isolates exposed to the same environment can exhibit phenotypic variations[279]. I removed these outliers for subsequent analyses in line with technical norms and because my work is exploratory, aiming to identify candidates for further investigation that may explain differences in pathogenicity.

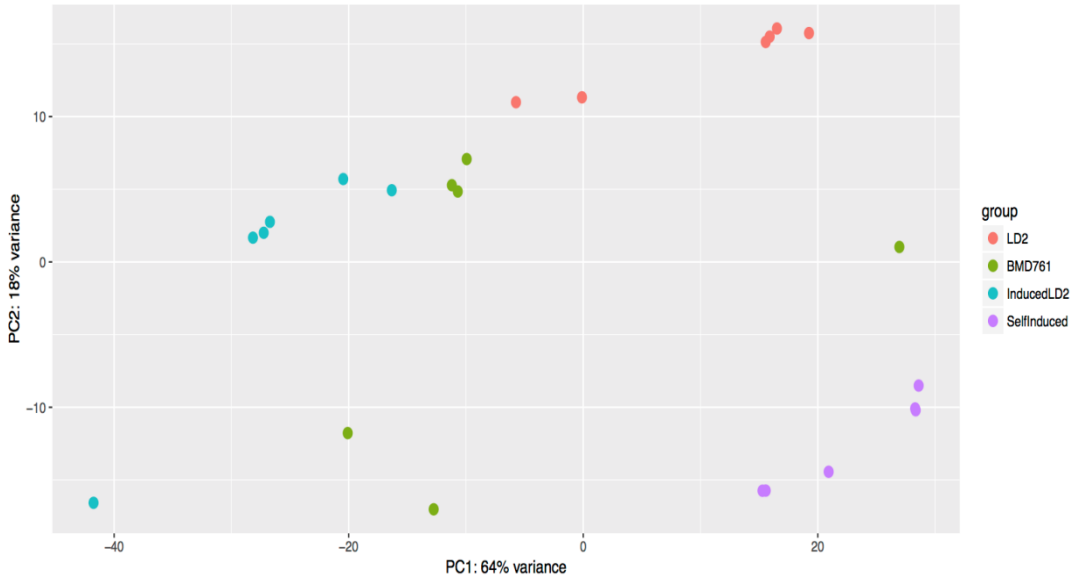


Figure 5-11 Principle component analysis of transcriptional profiles of 24 replicates. 4 isolates with 6 biological replicates each are included: environmental LD2, clinical BMD761 (derived from HIV uninfected patient), Induced LD2 (grown in culture filtrate from BMD761) and Self-induced LD2 (grown in culture filtrate from naive LD2).

After removal of 6 outliers, clustering was maintained by experimental group (Figure 5-12). PC1 increased in effect size explaining 72% of the variance in transcriptome profiles between samples. Replicates were grouped by experimental state. Strikingly, replicates of induced LD2 and BMD761 clustered closely to each other and separate from other isolates (naive and self-induced LD2), suggesting they had similar transcriptome profiles.

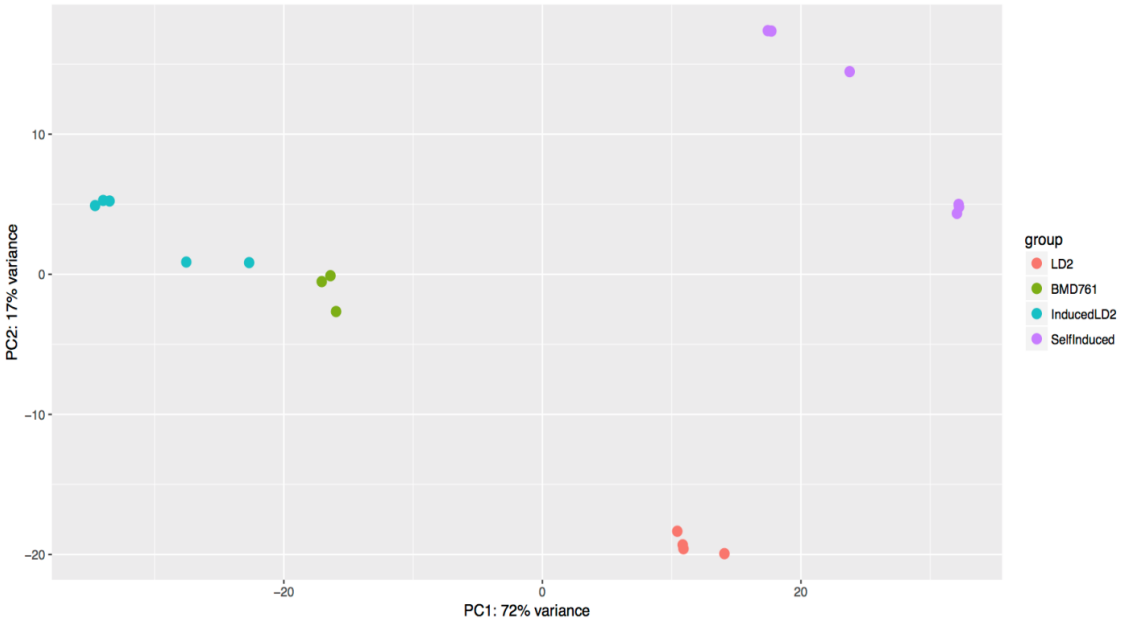


Figure 5-12 Principle component analysis of transcriptional profiles of 18 replicates after removal of 6 outliers. Replicates are separated by their corresponding isolates

5.4.7 Comparative analysis of transcriptome of clinical BMD761, induced LD2, self-induced LD2 relative to naive LD2 strains

Comparative analysis of 4 isolates (BMD761, Induced LD2, Self-Induced LD2 and LD2) was performed to test for differentially expressed genes. A total of 1604 genes were differentially expressed between BMD761, induced LD2, and self-induced LD2 relative to naive LD2 (Wald test, Benjamini-Hochberg adjusted P value < 0.05, log2 fold change ≥ 1 and ≤ -1). A hierarchical clustering of 1604 differentially expressed genes based on normalised read count per gene reveals two isolate-specific clusters (Figure 5-13). The transcriptional patterns of BMD761 and the induced LD2 cluster were similar and different from that of self-induced LD2 and naive LD2, with 1002 genes induced and 602 genes repressed in BMD761, induced LD2 and self-induced LD2 relative to LD2. The clustering of BMD761 and Induced LD2 in the heat map is consistent with the PCA plot, which also shows the proximity of BMD761 to Induced LD2 replicates.

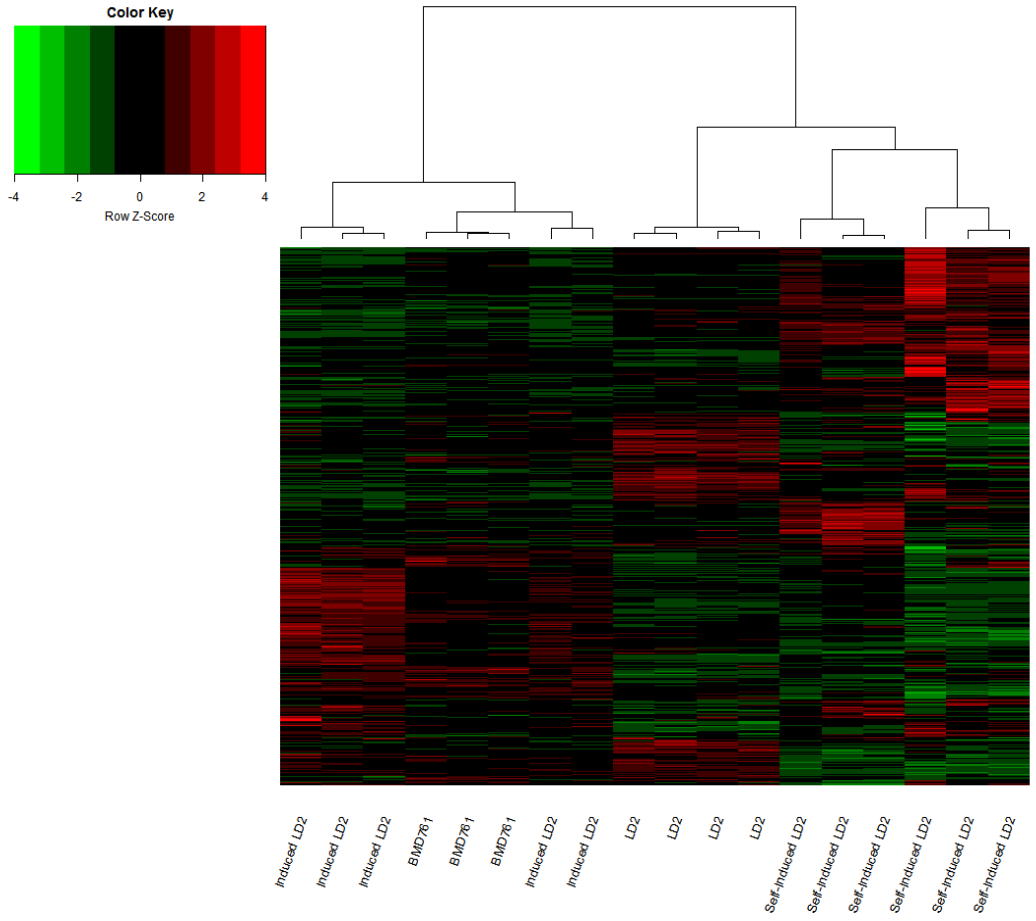


Figure 5-13 Hierarchical cluster analysis of gene expression based on normalised read counts per gene. The heat map displays expression patterns for 14 samples. Each sample corresponds to one column and each single gene to a single row. The colour key ranges from red for up-regulated gene to green for down-regulated genes across samples.

Pairwise comparisons of differentially expressed genes (Wald test, Benjamini-Hochberg adjusted P value <0.05, log2 fold change ≥1 and ≤-1) were made between four isolate groups (Figure 5-14). BMD761 and Induced LD2 had highly similar expression profiles, exhibiting the fewest number of differentially expressed genes, indicating they had similar transcriptome profiles and were different to the other experimental states. In contrast, Induced LD2 and Self-Induced LD2 had the least similarity, reflecting the role of Self-induced LD2 strain as control of induction effect.

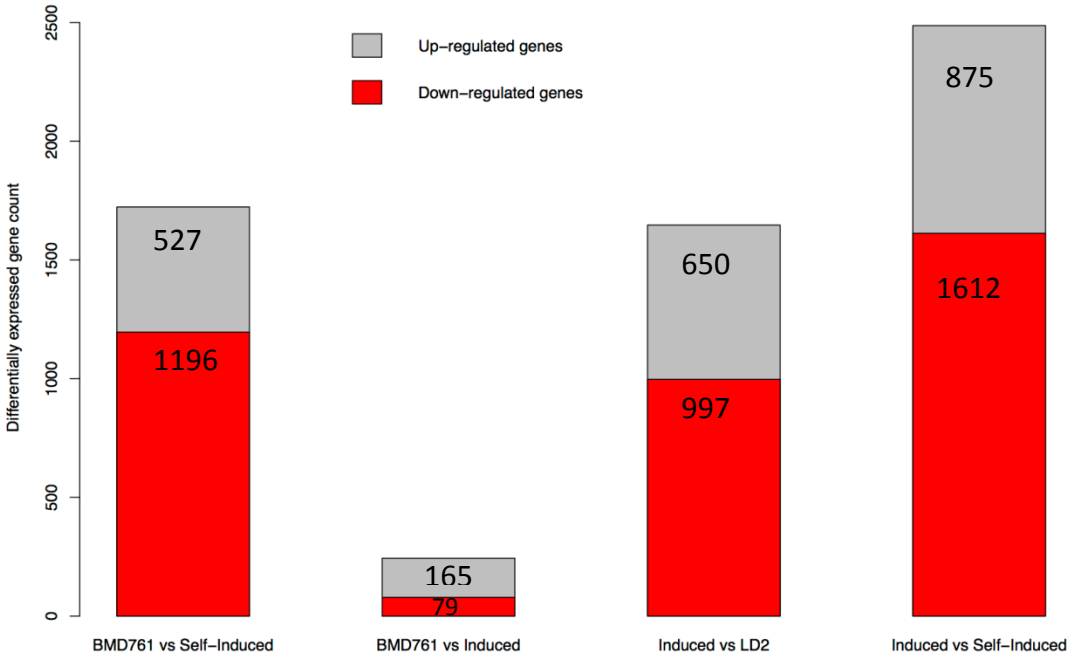


Figure 5-14 Pairwise comparison of differentially expressed genes for 4 isolates groups: BMD761 relative to Self-induced LD2 or Induced LD2, Induced LD2 relative to LD2 or Self-induced LD2 (Wald test, Benjamini-Hochberg adjusted P value < 0.05, log2 fold change $\geq$ 1 and $\leq$ -1). Numbers in stacked bars represent up- or down-regulated genes for pairwise comparison in each stacked bar.

5.4.8 Functional analysis using gene ontology

5.4.8.1 Biological process terms relevant to metabolism and cellular biosynthesis were enriched in BMD761 and Induced LD2 strains relative to naive LD2

Term ID	Description	Frequency (%)	log10 p-value
GO:0046128	purine ribonucleoside metabolic process	2.1	-2.9
GO:0009205	**purine ribonucleoside triphosphate metabolic process**	**1.7**	**-1.8**
GO:0009150	purine ribonucleotide metabolic process	2.3	-2.8
GO:0034645	**cellular macromolecule biosynthetic process**	**23.4**	**-2.3**
GO:0007049	cell cycle	11.6	-1.8
GO:0006310	**DNA recombination**	**3.8**	**-2.3**
GO:0006259	DNA metabolic process	8.2	-2.8

Table 5-13 Enriched biological process–associated GO terms of exclusively upregulated genes in BMD761 and Induced LD2 relative to naive LD2

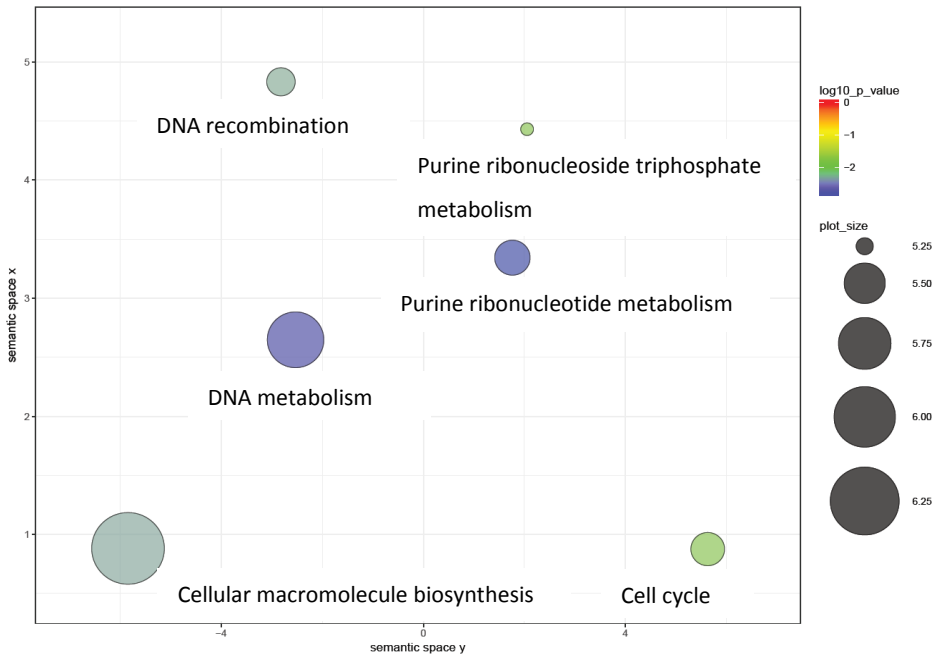


Figure 5-15 Biological process–associated GO terms of upregulated genes in BMD761/Induced LD2 relative to naive LD2. Bubble size represents frequency of GO terms in the GO annotation database (UniProt). Colour indicates significant level. Bubble proximity implies their semantic similarity

Of 1604 genes which were differentially expressed between BMD761, induced LD2, and self-induced LD2 relative to LD2, I curated 384 genes which were upregulated, and 349 genes which were down-regulated exclusively in BMD761 and induced LD2 (Wald test, Benjamini-Hochberg adjusted P value < 0.05, log2 fold change ≥ 1 and ≤ -1). The differentially expressed genes that were common to BMD761, Induced LD2 and Self-induced LD2 relative to LD2 were likely to be 'background' gene expression, perhaps driven by relative nutrient deprivation (since they were grown in only 75% fresh medium). I excluded this background 'noise' from the analysis. A list of significantly enriched GO terms of biological processes was summarised and visualised using semantic similarity REVIGO scatterplot (Figure 5-15 & Table 5-13).

GO terms related to cell metabolism were significantly enriched: purine ribonucleotide metabolism and DNA metabolism Moreover, upregulated genes involved in biosynthesis and cell cycle were frequently enriched including cellular macromolecule biosynthesis, and cell cycle (Table 5-13). Cell cycle involves separation of genetic material equally to daughter cells followed by cell division.

5.4.8.2 Molecular function GO terms related to protein synthesis and cell division were enriched in BMD761 and Induced LD2

Figure 5-16 displays enriched GO terms associated with molecular function. The most significantly enriched GO term was related to protein synthesis machinery: structural constituent of ribosome (Table 5-14). Other GO terms relevant to protein synthesis includes rRNA binding and translation elongation factor activity. Moreover, significantly enriched GO terms involved in cell division were identified: microtubule motor activity and microtubule binding.

I also analysed enrichment of GO terms for the cellular component domain. Only one GO term involved in cell division was significantly enriched: GO:0000922 spindle pole.

Term ID	Description	Frequency	log10 p-value
GO:0003735	structural constituent of ribosome	0.033	-30
GO:0003777	**microtubule motor activity**	**0.001**	**-5.57**
GO:0008017	microtubule binding	0.006	-9.05
GO:0005525	**GTP binding**	**0.015**	**-2.17**
GO:0004672	protein kinase activity	0.021	-1.56
GO:0003746	**translation elongation factor activity**	**0.002**	**-1.96**
GO:0019843	rRNA binding	0.010	-1.33

Table 5-14 Enriched molecular function GO terms of upregulated genes in BMD761/Induced LD2.

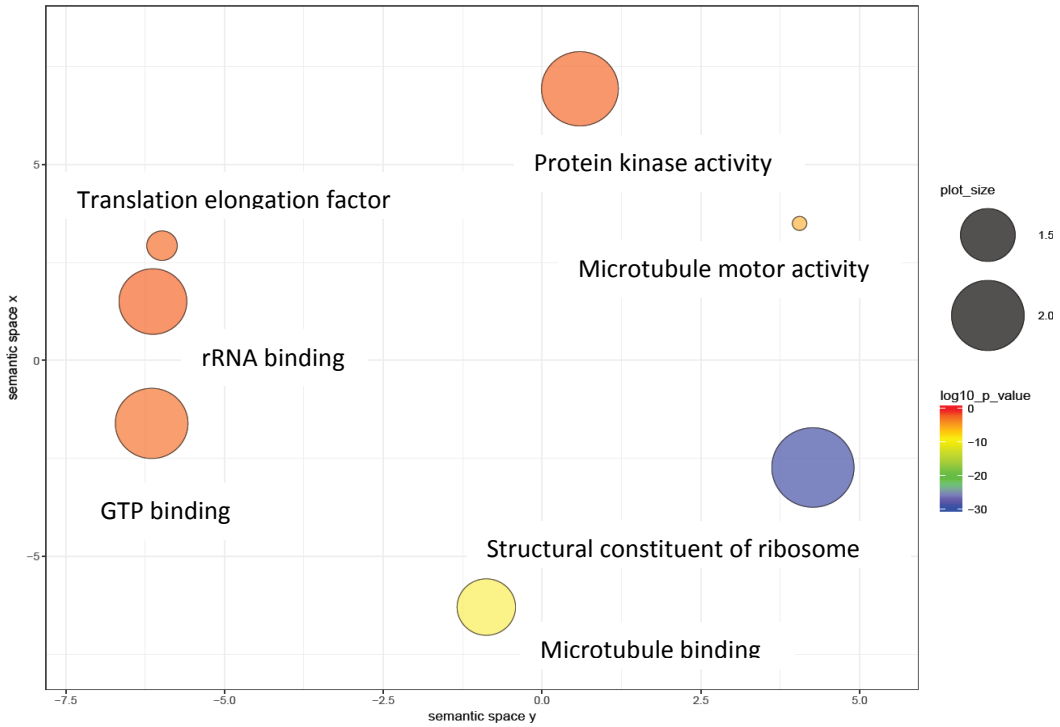


Figure 5-16 Molecular function–associated GO terms of upregulated genes in BMD761/Induced LD2. Bubble size represents frequency of GO terms in the GO annotation database (UniProt). Colour indicates significant level. Bubble proximity implies their semantic similarity

5.4.8.3 Gene ontology analysis of down-regulated genes in BMD761 and Induced LD2 strains relative to naive LD2

Only one GO term related to biological process was enriched in naive LD2 strains: oxidation-reduction process.

For GO terms associated with molecular function, genes relevant to oxidation-reduction were also enriched in naive LD2: oxidoreductase activity and protein disulfide oxidoreductase activity. Other GO terms include transporter activity: amino acid transmembrane transporter activity and substrate-specific transmembrane transporter activity (Table 15 & Figure 5-17).

Term ID	Description	Frequency	log10 p-value
GO:0005215	transporter activity	0.071	-2.21
GO:0015171	**amino acid transmembrane transporter activity**	**0.007**	**-1.46**
GO:0016831	carboxy-lyase activity	0.004	-2.25
GO:0015035	**protein disulfide oxidoreductase activity**	**0.003**	**-1.99**
GO:0016810	hydrolase activity, acting on carbon-nitrogen (but not peptide) bonds	0.013	-1.51
GO:0016616	**oxidoreductase activity**	**0.012**	**-1.60**
GO:0022891	substrate-specific transmembrane transporter activity	0.051	-2.51

Table 5-15 Enriched molecular function GO terms of downregulated genes in BMD761 and Induced LD2.

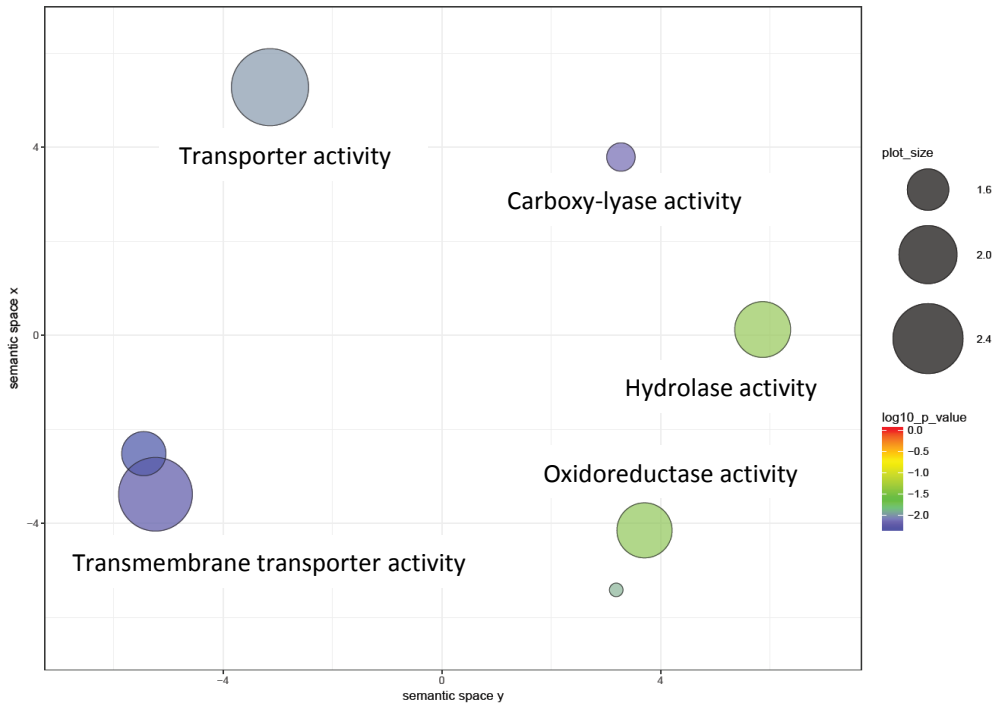


Figure 5-17 Molecular function–associated GO terms of down-regulated genes in BMD761 and Induced LD2. Bubble size represents frequency of GO terms in the GO annotation database (UniProt). Colour indicates significant level. Bubble proximity implies their semantic similarity

5.4.9 Differentially expressed genes associated with virulence in BMD761 and Induced LD2 relative to naive LD2

I screened 733 genes exclusively differentially regulated in BMD761 and Induced LD2 strains for virulence potential by excluding genes commonly regulated in Self-induced LD2 strains relative to naive LD2 strains.

5.4.9.1 Regulation of kinase genes in BMD761 and Induced LD2

Four out of five kinase genes were up-regulated in BMD761 and induced LD2 (Table 5-16)

GeneID	Product Description	Name	log2FoldChange
CNAG_03024	AGC protein kinase	RIM15	-1.0
CNAG_03184	**BUB protein kinase**	**BUB1**	**3.8**
CNAG_02675	CAMK/CAMKL/GIN4 protein kinase	HSL101	3.6
CNAG_06697	**TTK protein kinase**	**MPS1**	**1.4**
CNAG_03369	Wee protein kinase	SWE102	3.0

Table 5-16 Regulation of virulence-related kinase genes in BMD761 and Induced LD2 strains relative to naive LD2

5.4.9.2 Regulation of GAT201 target genes in BMD761 and Induced LD2 relative to LD2

Table 5-17 displays Gat201-bound genes that were differentially regulated in BMD761/Induced LD2 relative to naive LD2. Of 15 Gat201 target genes, only one was upregulated in BMD761/Induced LD2 strains.

GeneID	Product Description	Name	log2FoldChange
CNAG_06389	hypothetical protein	null	-3.4
CNAG_06493	**hypothetical protein**	**CNN2220**	**-2.7**
CNAG_04735	metalloproteinase	MEP1	-2.5
CNAG_00331	**anon-23da protein**	**YMR210W 01**	**-2.4**
CNAG_06186	hypothetical protein	CNM1740	-2.2
CNAG_00456	**hypothetical protein**	**null**	**-2.0**
CNAG_05147	hypothetical protein	CNI2090	-1.8
CNAG_00374	**hypothetical protein**	**null**	**-1.5**
CNAG_01601	lipase	ATG15	-1.5
CNAG_00165	**methylthioadenosine phosphorylase**	**MEU1**	**-1.5**
CNAG_03122	hypothetical protein	null	-1.2
CNAG_04737	**hypothetical protein**	**CNJ1790**	**-1.2**
CNAG_03848	glutathione transferase	CNB2310-B	-1.1
CNAG_05229	**stomatin family protein**	**CNI2770**	**-1.0**
CNAG_04517	hypothetical protein	null	2.9

Table 5-17 Gat201 target genes differentially expressed in BMD761 and Induced LD2 strains relative to naive LD2

5.4.9.3 Regulation of genes associated with murine lung infectivity in BMD761 and Induced LD2 relative to naive LD2

Mutations in a number of genes have been associated virulence attenuation in a mouse lung infectivity model[226]. Infectivity in this context refers to ability of pathogens to proliferate within mouse lung tissue. Table 5-18 displays genes previously associated with defects in *in vivo* lung infectivity, which are potential virulence factors. I found only one of these was induced in BMD761 and Induced LD2 strains.

GeneID	Product Description	Name	log2FoldChange
CNAG_01611	infection related protein of unknown function	LIV8	-2.0
CNAG_03894	**Binuclear cluster DNA-binding regulator**	**PDR802**	**-1.6**
CNAG_04611	ubiquitin-conjugating enzyme E2 H	UBC8	-1.0
CNAG_00181	**macrophage erythroblast attacher isoform 1**	**FYV10**	**-1.0**
CNAG_02930	virulence related protein of unknown function	LIV10	1.2

Table 5-18 Regulation of genes associated with murine lung infectivity in BMD761 /Induced LD2 strains relative to naive LD2.

5.4.9.4 Regulation of transcription factors in BMD761/Induced LD2 relative to naive LD2

Gene ID	Product Description	Name	log2FoldChange
CNAG_01438	cell cycle box factor subunit	SWI6	1.6
CNAG_03894	**transcription factor**	**PDR802**	**-1.6**
CNAG_04036	transcription factor	HSF3	-1.1
CNAG_04457	**transcription factor**	**FZC30**	**-1.1**

Table 5-19 Differential expression of transcription factors in BMD761 and Induced LD2 relative to naive LD2

Table 5-19 displays differential expression of gene encoding transcription factors (TF) of BMD761 and Induced LD2 relative to naive LD2. Mutants of these TF were attenuated in *G. mellonella* infection model and murine inhalation model [34], suggesting their potential virulence factors. Only one of them was up-regulated in BMD761/Induced LD2 strains: cell cycle box factor subunit *SWI6.*

5.4.9.5 Multiple established virulence-associated genes were upregulated in BMD761 and Induced LD2

I retrieved a list of 186 genes encoding established virulence factors (available at http://www.phidias.us/victors/index.php). Of which, 22 genes were differentially regulated in BMD761 and Induced LD2 strains relative to naive LD2 with 17 genes being upregulated (Table 5-20).

GeneID	Product Description	Name	log2FoldChange
CNAG_05472	meiotic recombination protein SPO11	SPO11	-2.1
CNAG_06574	**secreted antiphagocytic protein**	**APP1**	**-2.1**
CNAG_05015	catalase 4	CAT4	-1.7
CNAG_04217	**phosphoenolpyruvate carboxykinase (ATP)**	**PCK1**	**-1.3**
CNAG_05181	white-collar transcription factor; blue-light photoresponsive gene	BWC1	-1.0
CNAG_06500	**arginine metabolism transcriptional control protein**	**ARG1**	**1.0**
CNAG_02885	capsule-associated protein	CAP64	1.1
CNAG_02729	**hypothetical protein**	**FRR1**	**1.1**
CNAG_03263	translation elongation factor Tu	TUF1	1.2
CNAG_02531	**calcium-dependent protein kinase 2**	**CPK2**	**1.2**
CNAG_05998	rho-GTPase	RAC2	1.4
CNAG_01172	**parallel beta-helix repeat protein**	**PBX1**	**1.5**
CNAG_05274	STE/STE20/YSK protein kinase	null	1.6
CNAG_00418	**S-adenosylmethionine synthase**	**ILV2**	**1.6**
CNAG_02036	putative sugar transporter	CAS4	1.7
CNAG_01890	**methionine synthase**	**MET6**	**1.7**
CNAG_05968	rho-like GTPase	CDC420	1.7
CNAG_06231	**large subunit ribosomal protein L13**	**null**	**2.1**
CNAG_00779	large subunit ribosomal protein L27e	RPL27	2.4
CNAG_06487	**putative chitin synthase**	**CHS6**	**4.6**
CNAG_01907	PLK/PLK1 protein kinase	null	5.3
CNAG_05565	**hypothetical protein**	**null**	**5.7**

Table 5-20 Regulation of virulence-associated genes in BMD761 and Induced LD2 strains relative to naive LD2**.** Null means there are no function described or no ortholog for these genes.

5.4.9.6 Regulation of transporter genes in BMD761 and Induced LD2 relative to naive LD2

Transporters play important roles in yeast pathobiology[262]. For example, major facilitator superfamily (MFS) transporters are responsible for multiple drug resistance to fungus[276] [277]. Others enable cryptococci adaptation to different ecological niches, such as pigeon excreta, murine lung and human CSF [259] [278] [201].

Table 5-21 shows up-regulation of gene encoding transporters as many as down-regulated transporter genes in BMD761 and Induced LD2 strains.

Interestingly, multiple genes encoding sugar, zinc ion, mannose and monosaccharide transporter were exclusively enriched in BMD761 and Induced LD2 (Table 5-21). Of upregulated transporters, the zinc ion transporter (CNAG_00895) was the most highly induced.

Gene ID	Product Description	Name	log2FoldChange
CNAG_05662	sugar transporter	PTP1	-2.0
CNAG_07902	**AAT family amino acid transporter**	**null**	**-1.6**
CNAG_04784	monosaccharide transporter	null	-1.4
CNAG_04617	**OPT family small oligopeptide transporter**	**null**	**-1.3**
CNAG_07874	sugar transporter	null	-1.1
CNAG_06259	**MFS transporter, SP family, general alpha glucoside:H+ symporter**	**null**	**-1.1**
CNAG_01055	phospholipid transporter	null	-1.0
CNAG_06119	**spermine transporter**	**null**	**-1.0**
CNAG_02600	tartrate transporter	null	1.1
CNAG_05254	**UDP-galactose transporter**	**null**	**1.2**
CNAG_06545	pim1 protein RNA transporter 2	null	1.2
CNAG_01683	**putative monosaccharide transporter**	**STL1**	**1.3**
CNAG_02036	putative sugar transporter	CAS4	1.7
CNAG_03426	**GDP-mannose transporter 2**	**GMT2**	**1.9**
CNAG_06963	sugar transporter	null	2.3
CNAG_00895	**zinc ion transporter**	**ZIP1**	**3.3**

Table 5-21 Differential regulation of transporter genes in BMD761 and Induced LD2 strains relative to naive LD2

5.4.9.7 Regulation of hypothetical proteins in BMD761 and Induced LD2 relative to naive LD2

Of 733 differentially expressed genes, a third of them (263) are hypothetical proteins. I curated a list of hypothetical proteins which were highly upregulated (log2FoldChange greater than 3) and highly down-regulated (log2FoldChange less than -3 in BMD761 and Induced LD2 strains (Table 5-22 & 5-23)

GeneID	Name	Predicted function	Ortholog	log2FoldChange
CNAG_05565	null	null	null	5.7
CNAG_05853	null	protein kinase activator	C. gattii	4.9
CNAG_03539	null	SprT family metallopeptidase	Aspergillus clavatus	4.7
CNAG_00110	HPI3	Rho GTPase binding, and role in actin cytoskeleton organization	Aspergillus fumigatus	4.6
CNAG_01389	null	null	null	4.4
CNAG_06902	null	ENTH-domain containing protein	Phanerochaete chrysosporium	4.4
CNAG_05863	null	null	null	4.0
CNAG_05576	null	null	null	4.0
CNAG_00406	null	null	null	3.9
CNAG_02724	null	null	null	3.6
CNAG_01334	null	TAT-binding protein-like protein 7	C. gattii	3.6
CNAG_04227	null	null	null	3.5
CNAG_02060	null	null	null	3.4
CNAG_05699	null	null	null	3.4
CNAG_06795	null	Chitinase	Mucor circinelloides	3.4
CNAG_04633	null	null	null	3.4
CNAG_03125	null	ER to Golgi transport-related protein	C. gattii	3.4
CNAG_01903	null	null	null	3.4
CNAG_02235	null	N2-dimethylguanosine tRNA methyltransferase	Phycomyces blakesleeanus	3.3
CNAG_03037	null	microtubule anchoring at spindle pole body, mitotic sister chromatid segregation	Aspergillus fumigatus	3.2
CNAG_01105	null	null	null	3.0
CNAG_04963	null	null	null	3.0

Table 5-22 Hypothetical proteins up-regulated in BMD761 and Induced LD2 isolates

GeneID	Predicted function	Ortholog	log2FoldChange
CNAG_02118	null	null	-4.0
CNAG_06389	null	null	-3.4
CNAG_00679	null	null	-3.4
CNAG_01056	cytoplasm localization	Aspergillus fumigatus	-3.4
CNAG_00588	null	null	-3.1

Table 5-23 Hypothetical proteins down-regulated in BMD761 and Induced LD2. Null means there are no function described or no ortholog for these genes.

5.5 Discussion

5.5.1 Enrichment of biological process terms relevant to cellular biosynthesis, metabolism and cell cycle may be attributed to enhanced virulence of clinical isolates and induced LD2 strain.

Four out of 7 biological process GO terms were common between clinical isolates and Induced LD2/BMD761 relative to naive LD2 (Table 5-3 and 5-13). Enrichment of GO terms related to metabolism (purine ribonucleotide metabolism, coenzyme metabolism, carbohydrate metabolism and DNA metabolism) indicates more intense metabolic activity in clinical isolates (Table 5-3). Biosynthesis is more active in clinical isolates with enrichment of genes annotated to biosynthesis (cellular macromolecule biosynthesis, organic substance biosynthesis, small molecule biosynthesis and cellular biosynthesis). Upregulation of genes annotated to GO terms for synthesis of protein and DNA, including structural constituent of ribosome, translation elongation factor activity and DNA-directed DNA polymerase activity, is likely to contribute to cellular macromolecule biosynthesis.

These induced activities (metabolism & biosynthesis) may explain for significantly higher burden of clinical isolates compared to environmental isolates (Figure 4-19). My finding is consistent with a study comparing transcriptome profiles between more active *C. neoformans* cultured in YPD and *C. neoformans* with less active growth in *ex vivo* CSF. Chen *et al* found enrichment of GO terms related to cellular biosynthesis and structural constituent of ribosome of active yeasts grown in YPD[262].

Steen *et al* used serial analysis of gene expression (SAGE) to compare transcription profiles of cells isolated in rabbit CSF and cells grown *in vitro.* They also identified abundance of sequences annotated to GO terms relevant to protein biosynthesis, carbohydrate metabolism, and DNA metabolism for cells grown *in vivo*[255].

Ferrareze *et al* compared transcriptome of *C. gattii* isolated in murine BAL (bronchoalveolar lavage) and those cultured *in vitro* confirmed enrichment of biological process terms relevant to cellular biosynthesis and metabolism in the former culture condition[280].

Taken together, these findings imply that *C. neoformans* accelerate intracellular growth in clinical settings by speeding up biosynthetic and metabolic processes.

Moreover, I found upregulation of genes annotated to GO terms relevant to cell cycle in Induced LD2/BMD761 relative to naive LD2 (Table 5-13). Cell cycle is the process of cell duplication to produce two genetically identical daughter cells. There is evidence that throughout the cell cycle, 20% of all genes are periodically

transcribed. Moreover, DNA replication genes, mitosis genes and 40 genes associated with virulence are frequently co-expressed, suggesting a connection between cell cycle regulators and virulence pathways[281 282]. Genes involved in cell cycles are potential drug targets by affecting cell growth and virulence factors coordinated with such genes. Thus cell proliferation and greater activities of metabolism and biosynthesis can explain for the higher fungal burden of BMD761/Induced LD2 strains than naive LD2.

Only two GO terms for biological process were enriched in environmental isolates, one of which was purine ribonucleotide biosynthesis (Table 5-5). This supports a hypothesis that biosynthetic activities in environmental isolates were much less active than those in clinical isolates. Enhanced virulence of Induced LD2 strains could be attributed to upregulation of genes involved in cellular biosynthesis and metabolism as well as cell cycle.

In contrast to enrichment of multiple GO terms in BMD761 and Induced LD2, only one GO term for biological process was enriched in naive LD2: oxidation-reduction process. I am not clear why this process was induced in environmental isolates.

5.5.2 Enrichment of molecular function terms relevant to cell division and peptide synthesis may be attributed to increased virulence of clinical isolates and BMD761/induced LD2 strains.

Six out of 8 molecular function-associated GO terms was common between clinical isolates and Induced LD2/BMD761 (Table 5-4 and 5-14). They are involved in cell division and peptide synthesis (translation elongation factor activity, structural constituent of ribosome).

Enrichment of molecular function GO terms in clinical isolates and BMD761/induced LD2 are related to cell division via microtubule activities: microtubule motor activity and microtubule binding (Table 5- 5). Microtubules are tube-like structure consisting of tubulin protein. Apart from composition of cytoskeleton, microtubules are integral to multiple cellular processes in eukaryote: transport of organelles through the cytoplasm, cell motility and cell division (separation of chromosome during mitosis and meiosis)[283]. Moreover, cellular component GO term exclusively enriched in clinical isolates and BMD761/induced LD2 is relevant to cell division too: spindle

pole, which is composed of microtubules. Each pole is involved in portioning chromosomes evenly to each daughter cell.

Besides, I found upregulation of genes annotated to GO terms associated with rRNA binding in Induced LD2/BMD761. rRNA binding occur during formation of ribonucleo-protein complex for ribosome. Taken together, these activities may contribute to robust protein biosynthesis and proliferation of BMD761/Induced LD2, thereby promoting their faster growth inside *G. mellonella*.

5.5.3 *GAT201* target genes are dispensable for virulence of clinical isolates and BMD761/Induced LD2 strains

I found 19 of *GAT201* target genes were differentially regulated in my clinical and environmental *C. neoformans* strains in YPD culture broth (Table 5-6). The product of one of these, Qsp1, was previously recognized to be an autoregulatory signaling factor[246]. The *QSP1* gene encodes a peptide precursor consisting of a signal sequence and a 24-amino acid pro-peptide (proQsp1). proQsp1 is synthesized intracellularly. Once exported extracellularly, it matures through cleavage mediated by the cell-associated protease Pqp1. The mature Qsp1 enters surrounding cells via peptide transporter Opt1, where it initiates morphological and virulence changes. *QSP1*-deficient mutants exhibit attenuated virulence in the murine inhalational infection model[246]. The downregulation of *QSP1* in my clinical isolates may be compensated by upregulation of other Gat201 target genes which makes clinical isolates significantly more virulent. The other 18 genes may be dispensable for virulence as their mutants did not show remarkable defects in phagocytosis by macrophage[116]. The limitation of the referenced study by Chun *et al* is that the H99 strain and its mutants were used and cultured in DMEM (tissue culture broth) while I used clinical strains of ST4 and ST5 and cultured them in YPD broth.

Of 19 *GAT201* target genes in Table 5-6, six were also down-regulated in BMD761/Induced LD2 (Table 5-17) .However, their mutants were not attenuated in macrophage infection model[116], suggesting their dispensable role for virulence in clinical isolates and BMD761/Induced LD2 strains.

5.5.4 *LIV10*, gene associated with murine lung infectivity, is responsible for enhanced virulence of clinical isolates and BMD761/Induced LD2

I found 6 differentially expressed genes in clinical and environmental isolates (Table 5-7), whose mutants are known to exhibit defects in lung infectivity but have normal for growth at 37^{o}C, capsule enlargement and melanin production[284 285]. I found upregulation of *LIV10* gene in clinical isolates and BMD761/Induced LD2. This suggests contribution of this gene to enhanced virulence of such isolates in *G. mellonella* model.

However, I detected down-regulation of *LIV5* and *ENA1* in clinical isolates. The *LIV5*–deficient strain has exhibited increased survival in murine inhalation model[226]. This gene is homologous to one gene required for pathogenicity in the plant fungal pathogen *Magnaporthe grisea*[226]. *ENA1* encodes a potassium/sodium ATPase transporter[201]. This gene is required for *C. neoformans* survival in *ex vivo* CSF. *ENA1* mutants fail to grow in CSF after 4 days of incubation. The mutant strain also displayed virulence attenuation in rabbit meningitis model [201]. Down-regulation of *LIV5* and *ENA1* in clinical isolates suggests that they work independently and may be compensated by other virulence factors. Likewise, I found down-regulation of *PDR802* (Table 5-18) in BMD761/Induced LD2 vs naive LD2. The DNA-binding regulator *PDR802*–deleted mutants have displayed growth defect in mouse lung compared to the wild-type strain, suggesting their role in pathogenicity[285]. The repression of this gene in BMD761 and Induced LD2 implies their dispensable role for virulence in *G. mellonella*.

5.5.5 Four kinase genes contribute to increased virulence of clinical isolates and BMD761/Induced LD2

Kinases play central roles in metabolic pathways and signaling cascades in *C. neoformans*[227 140 155], thereby regulating its pathogenicity. For example, the protein kinase A catalytic subunit (Pka1) activate downstream proteins in the cAMP-PKA signaling pathway, governing cell wall synthesis, expression of transmembrane transporters, melanin formation, and stress response of *C. neoformans*[140 155]. Interestingly, I detected elevated expression of seven kinase genes in clinical isolates

(Table 5-8). Four of these kinases (*BUB1*, *HSL101*, *MPS1* and *SWE102*) were commonly upregulated in clinical isolates and BMD761/Induced LD2 (Table 5-16). These kinase mutants have been related to pathogenicity using the *G. mellonella* infection model and murine infectivity assay[227], suggesting their potential drug targets. The elevated expression of kinase genes suggests their involvement in virulence of clinical isolates and BMD761/Induced LD2 strains.

5.5.6 Transcription factor role in pathogenicity

Transcription factors govern general biology and pathogenicity of *C. neoformans* through activation or repression of effector proteins in signal transduction pathways [34]. I identified 6 differentially expressed genes which have been previously confirmed to regulate virulence of *C. neoformans* using the murine or insect infection model[34]. Two genes of particular note were *PDR802* and *GAT201* (Table 5-8), whose mutants exhibited reduced virulence in such models. I found down-regulation of *GAT201* in clinical isolates. This can be associated with the down-regulation of *GAT201* target genes that I mentioned in table 5-6. The downregulation of *GAT201* may be compensated by upregulation of other transcription factors in clinical isolates which made clinical isolates significantly more virulent than environmental isolates. I detected upregulation of zinc-finger transcription factor *PDR802*, which may confer enhanced virulence on clinical isolates.

Additionally, I identified upregulation of the cell cycle box factor subunit gene *SWI6* (Table 5-19*)* in BMD761/Induced LD2. This gene was reported to drive cell-cycle genes[281]. *SWI6* mutants cause growth defect in murine lung compared to wild type strains, implicating their potential virulence attribute[34]. The presence of this gene in BMD761/Induced LD2 suggests its contribution to enhanced virulence of these strains in *G. mellonella*.

5.5.7 BMD761/Induced LD2 strains displayed upregulation of multiple virulence-associated genes

Cryptococcal pathogenicity is determined by composite of virulence factors, which are coordinately regulated[286]. Established virulence factors are of interest due to potential drug targets. I found elevated expression of multiple virulence factors in clinical isolates compared with environmental isolates, and in induced LD2, lending biological plausibility to the increased virulence observed in the *Galleria* model and potentially explaining that increased virulence (Table 5-10). Of 19 induced genes in clinical isolates, 15 genes were expressed in both BMD761 and Induced LD2 (Table 5-21)

5.5.8 Cell wall and capsule-associated genes are upregulated in hypervirulent and induced strains

A number of genes relevant to synthesis of cell wall and capsule were upregulated in clinical isolates. These included chitin synthase (*CHS2, CHS6*) & capsule structure designer proteins (*CAS4, CAS32*). Chitin is indispensable to the integrity of the fungal cell wall. *C. neoformans* has 8 putative chitin synthases (*CHS1* to *CHS8*)[287]. They are periodically induced during the *C. neoformans* cell cycle[288]. *CAS* genes are important for capsule synthesis. *CAS* gene-deficient mutants have resulted in capsule defects[289]. Of all elevated genes in clinical isolates, *CHS6* was the most highly induced (Table 5-10). This gene was highly expressed in BMD761/Induced LD2 too (Table 5-21). Taken together, upregulation of these genes may account for significantly larger capsule in clinical isolates compared to environmental isolates in larvae (Figure 4-20).

Apart from genes upregulated in clinical isolates, I detected upregulation of *CAP64* gene in BMD761/Induced LD2. This gene is critical to capsule formation. Deletion of *CAP64* leads to loss of capsule as well as pathogenicity in mouse infection model[290]. Another gene involved in capsule synthesis is *PBX1*, parallel beta-helix repeat protein, participating in synthesising glucuronoxylomannan (GXM). GXM accounts for 88% of capsular mass. *PBX1* mutants exhibit clumpy cells and decreased capsule integrity, leading to virulence attenuation in mouse infection model[291]. I found

upregulation of *PBX1* in clinical isolates and BMD761/Induced LD2, suggesting its contribution to enhanced virulence of these isolates

The cAMP-PKA signal transduction pathway regulates capsule production and melanin formation. *PKA1* mutants are avirulent with defects in capsule size and melanin formation. Pka1 activity modulates up to 302 proteins of *C. neoformans*[292], including the large subunit ribosomal protein L27e encoded by *RPL27* gene, which was upregulated in clinical and BMD761/Induced LD2 strains

5.5.9 Oxidative stress-regulated genes

To manage the harsh environment in macrophages, *C. neoformans* elaborate antioxidant defense mechanisms to neutralise reactive oxygen species elicited by phagolysomes. The catalase gene family together with superoxide dimutase Sod1 & Sod2, cytochrome c peroxidase Ccp1, thioredoxin reductase Trr1, alternative oxidase Aox1 contribute to resistance against oxidative stress[293] [284]. I detected upregulation of *CAT4* in environmental isolates (Table 5-10). However, Cat4 has not demonstrated detectable catalase activity *in vitro* [293] although its transcript was detected by real time quantitative PCR. Moreover, mortality of mice infected with *CAT4* mutant was not significantly different to those infected with wild type strains. This suggests that *CAT4* is compensated by other *CAT* genes (*CAT1, CAT2* and *CAT3*) and other components of the antioxidant system.

Another oxidative stress-regulated genes was highly induced in environmental isolates was *CCP1*. This gene is involved in the stress-activated HOG signaling pathway[284] for adaptation to hostile host environment. The down-regulation of this gene in clinical isolates suggests that this gene works independently and can be compensated by other components of the antioxidant system.

5.5.10 Other virulence genes

I also found induction of methionine synthase (*MET6*) in clinical isolates. Methionine-cystein biosynthesis has been extensively investigated in a number of fungal pathogens due to their relevance as a potential drug target. *MET*-deficient mutants lose viability due to methionine starvation. They are avirulent in the murine inhalation model[294]. Upregulation of *MET6* will lead to expression of methionine.

This explains the enrichment of GO terms relevant to cellular biosynthesis in clinical isolates I mentioned above.

In this study I detected two-fold upregulation of carbonic anhydrase gene (*CAN2*) in clinical isolates. During growth in natural habitats, such as tree hollows, soil, and avian excreta, *C. neoformans* are exposed to low ambient CO_2 levels (0.033 %). However in mammalian hosts the CO_2 concentration can rise to 5 %; *C. neoformans* adapts to such elevated levels of CO_2 by inducing this enzyme to efficiently convert CO_2 to bicarbonate (HCO_3^-)[295]. Bicarbonate in turn stimulates Adenylyl Cyclase Cac1 to regulate biosynthesis of capsule through the GPA1-cAMP signaling pathway[139]. Capsule confers advantage on clinical isolates for survival in mammalian hosts.

Of down-regulated genes listed in Table 5-10 and Table 5-21, *APP1* is the most down-regulated in clinical isolates and BMD761/Induced LD2 strains. This gene encodes an antiphagocytic protein that is secreted to the extracellular space. However, in an impaired immunity mouse model (deficient for T and NK cells) the *APP1* mutant exhibited hypervirulence and dissemination with increased number of fungal cells recovered from the brain[113]. The down regulation of *APP1* in hypervirulent strains suggests that phagocytosis may infect be beneficial for the survival of these strains, perhaps through providing an environment for replication and dissemination.

In my study, expression of *BWC1* was elevated in environmental isolates. Light is known to regulate cryptococcal development through inhibition of mating mediated by the gene *BWC1*[296]. In addition, the *BWC1* gene protects cells from UV radiation. Its disruption results in hypersensitivity to ultraviolet light compared with wild type strains. However, it likely has a role in virulence – mutants also appear significantly less pathogenic in murine inhalation model than the wild type strain. Down-regulation of BWC1 in clinical isolates suggests that any role in virulence is secondary to other functions.

Another gene exhibited elevated expression in clinical isolates and BMD761/Induced LD2 is the translation elongation factor *TUF1* responsible for mitochondrial protein synthesis. Production of mitochondrial protein is indispensable for ATP generation through oxidative phosphorylation[286].

The calcium-dependent protein kinase 2 (*CPK2*) was upregulated in BMD761/Induced LD2 strains. *CPK2*-deficient strains were involved in cell cycle regulation and attenuated in murine infectivity model[227]. The presence of this gene can explain for higher fungal burden of these strains compared to naive LD2 in *G. mellonella.*

I screened 19 virulence-associated genes, 9 of which were down-regulated in clinical isolates. This can be explainable by the fact that cryptococcal pathogenicity is related to virulence composite. Virulence factors (capsule size, melanin formation, phospholipase activity, protease activity) partially contribute to overall pathogenicity of *C. neoformans* [275] [274]. Litvintseva *et al* found clinical and environmental *C. neoformans* with identical MLST profile. They exhibited similarity of such virulence traits but differed in murine virulence[62]. Non-capsular *C. neoformans* can cause cryptococcal meningitis and pulmonary cryptococcosis in HIV patient and HIV-uninfected patients, respectively [297] [298]. Multiple genes have overlapping functions in response to genetic perturbation, so lacking of a virulence attribute can be compensated by others[288]. For instance, *C. neoformans* utilise redundant systems to neutralise superoxide radicals. Superoxide dimutase (SOD) detoxifies such radicals through the conversion of them into hydrogen peroxide and oxygen. However, *SOD1* mutants maintain virulence in the mouse model, suggesting other radical quenchers compensate for loss of *SOD1*. Such quenchers could include mannitol and melanin[299]

5.5.11 Upregulation of transporter genes involved in nutrient transportion and capsule components

Transporters were frequently identified in transcriptional analyses, implicating their role in regulating fungal physiology. The two superfamilies MFS transporters and sugar transporters were the most abundant in the *Cryptococcus* genome[259]. Overexpression of MFS transporter was discovered in Fluconazole-resistant clinical isolates[276]. In this study, I detected a number of transporters belonging to these two families that were differentially expressed in clinical versus environmental isolates. Notably, the expression of six sugar transporters and two monosaccharide transporter genes were elevated in clinical isolates. Expression of these genes may facilitate nutrient uptake of *C. neoformans* inside the phagosome. Moreover, I

detected enrichment of two genes encoding myo-inositol transporters belonging to the sugar transporter family. Inositol is abundant in human CSF (25 mg/L compared to an average of 4.3 mg/L in plasma)[300]. *C. neoformans* can utilize myo-inositol as a sole carbon source through the myo-inositol transporter gene family during their survival in environmental niches and the human host[300]. Therefore this may confer a growth advantage in CSF. Differential expression of myo-inositol transporter genes between clinical and environmental isolates could be validated by culturing these isolates in CSF known as *ex vivo* CSF survival testing.

I also detected up-regulation of the ATP-binding cassette transporter gene *PDR5* (Table 5-11). This gene is highly induced in response to osmotic stress. *PDR5* mutants are more susceptible to fluconazole than wild type strains, suggesting its potential role in efflux of antifungals for detoxification. In *S. cerevisiae PDR5* is responsible for cellular detoxification and has been associated with multiple drug resistance[284]. Upregulation of this gene may explain for fluconazole heteroresistance inherent to clinical isolates[301].

Two zinc transporters were upregulated in clinical isolates (Table 5-11) but either of them was expressed in BMD761/Induced LD2 (Table 5-21). Zinc is a vital micronutrient for all organisms. Many enzymes require zinc for catalysis. In yeast cells, zinc is required for optimal function of approximately 3% of the proteome [302]. In *C. gattii*, sibling species of *C. neoformans*, it has been shown that zinc deprivation leads to increase transcript levels of zinc transporters *ZIP1* and *ZIP2* [302]. Zinc is co-factor of carbonic anhydrase[303]. Expression of zinc transporters facilitate utilization of zinc for functional carbonic anhydrase, whose gene *CAN2* was also highly induced (Table 5-10).

Taken together, the elevated expression of nutrient-related transporters (sugar, zinc, iron, amino acid) likely represents a growth advantage within the host, and may account for the higher fungal burdens seen in *Galleria* when infected with clinical isolates (Figure 4-20).

Cryptococcal capsule is composed of glucuronoxylomannan (GXM) and galactoxylomannan, as well as mannoproteins. Mannose acts as backbone of GXM accounting for the majority of capsule mass. To produce manosylated compounds,

the nucleotide sugar GDP-mannose is required as a donor for enzymes to synthesize GXM. *C. neoformans* has at least two GDP-transporters to transport mannose donors to the Golgi apparatus: *GTM1* and *GTM2*. I detected upregulation of *GTM2* in clinical isolates and BMD761/Induced LD2. Under capsule-inducing condition, *GTM1* mutants have striking defects in capsule size, but capsule of *GTM2* mutants was similar to that of wild type strain[304]. However, this finding is based on serotype D strain JEC21 and reference strain H99. My finding suggests a potential role for *GTM2* in the virulence of *C. neoformans*.

I also found upregulation of the glycerol symporter *STL1* in BMD761/Induced LD2. This gene was induced in response to osmotic stress by transporting osmolytes to counteract external osmotic change.[284]

I detected down-regulation of polyol transporter protein 1 (*PTP1*) in BMD761/Induced LD2. This is responsible for cryptococcal growth in media containing 5- or 6- carbon polyols, such as mannitol, sorbitol and arabitol. Upon internalization by amoeba and murine macrophage, *PTP1* has been highly induced. However, the *PTP1* mutant has not exhibited defects in *in vitro* phenotypes implicated for virulence (growth at 37^{o}C, melanin production, phospholipase activity) and survival of this strain has not been significantly different from the wild type strains in mouse and *G. mellonella* infection model.[33] This supports the notion that transcriptional response to phagocytosis activates many genes but not all of them are essential for virulence. Thus down-regulation of *PTP1* may not affect virulence of BMD761 and Induced LD2 strains

5.5.12 Differential expression of hypothetical proteins

Interestingly, I found that upregulated hypothetical proteins outnumber down-regulated hypothetical proteins (22 vs 5) in BMD761 and Induced LD2. Some are potential metallopeptidases, proteases and kinases (Table 5-22 & 5-23). Others involve transport from the Endoplasmic Reticulum to the Golgi apparatus, endocytosis, cell wall synthesis and phosphorylation of signaling proteins. Notably, there was an upregulated gene (CNAG_02235) encoding one enzyme involving stabilization of tRNA structure for efficient protein synthesis (tRNA

methylatransferase) and another upregulated gene (CNAG_03037) responsible for cell division mediated by chromatid segregation.

5.6 Conclusion

I detected 1700 genes which were differentially expressed between clinical and environmental isolates. A number of GO terms relevant to metabolism and cellular biosynthesis were enriched in clinical isolates. Besides metabolism and cellular biosynthesis, enriched GO terms relevant to cell cycle and cell division can explain for enhanced virulence of BMD761/Induced LD2.

I found similarity of transcriptome profiles between clinical BMD761 and Induced LD2. This may explain the enhanced virulence of Induced LD2 compared to naive LD2. I also identified a number of genes which were previously confirmed as virulence factors that were upregulated in clinical isolates and BMD761/Induced LD2. A third of differentially expressed genes between BMD761/Induced LD2 and naive LD2 were hypothetical proteins (263 genes). This implies the importance of uncharacterized genes responsible for enhanced virulence of these strains.

Chapter 6

General discussion

6.1 Introduction

My research was driven by the compelling biological observation that it is a single lineage of *C. neoformans* (VNIa-5) that is responsible for the vast majority of cryptococcal disease in immunocompetent patients in Vietnam, whereas HIV infected patients are infected by a diverse population of *C. neoformans*.

The specific aims of my thesis were to:

1. To determine whether there is temporal-spatial clustering of *C. neoformans* by genetic lineage and HIV status in Vietnam,
2. To identify the ecological niche of particular lineages of *C. neoformans*,
3. To define the diversity of the Vietnamese *C. neoformans* population in the environment and prevalence of particular lineages,
4. To characterise the pathogenic phenotypes of the environmental isolates,
5. To compare the virulence between and within different genotypes and sources using the *G. mellonella* model,
6. To define the transcriptional differences associated with variability in virulence and
7. To identify gene candidates associated with variability in virulence between isolates.

First, I mapped clinical cases of meningitis and used clustering analysis in an attempt to identify the broad ecological niche of *C. neoformans* isolates of different lineages. To better understand the diversity of the *C. neoformans* population, their relative prevalence and ecological niches I then used randomized environmental sampling, supplemented by targeted sampling (because of low yield). It must be said in this aim I was not as successful as I hoped, although I have to some extent characterized the population distribution at the genus level. Using ITS barcoding for speciation I

identified that the species most commonly recovered are rarely (or not at all) pathogenic to humans.

Given the difficulty I had in isolating the well-recognized human pathogenic forms of *Cryptococcus* from the environment (*C. neoformans* or *C. gattii*), I decided to define the virulence-associated traits of the environmental isolates in an attempt to understand why, while they are prevalent, they are not causing disease in the local population. I described the thermotolerance, melanin production, capsule enlargement, *ex vivo* CSF survival and survival in the *Galleria* infection model.

I had developed the *Galleria* model to enable me to define differences between the virulence of clinical isolates of different lineages, because of the association of VNIa-5 with immunocompetence. I was able to show that this model of innate immunity could distinguish differences in virulence between VNIa-5 and other lineages. However, the use of the model led to the interesting observation of my thesis that isolates have an 'ecological imprint' – their source (environmental/HIV—infected patient/immunocompetent patient) determines the virulence phenotype in *Galleria*. However, the most exciting result of my thesis is the observation that there is intercellular signaling that can lead to upregulation of virulence of naive isolates by 'hyperpathogenic' isolates (by which I mean they can cause disease in immunocompetent patients). I then used RNAseq and validated the upregulated virulence phenotype by showing it was associated with upregulation of previously identified virulence associated genes, and identified a number of hypothetical proteins that are potential drivers of the induction phenomenon and virulence itself.

6.2 Summary of principal findings

1. I detected temporal-spatial clustering of cases of cryptococcal in Ho Chi Minh city and bordering well-industrialised provinces (Ba Ria-Vung Tau and Dong Nai). The temporal signal has to be interpreted in the face of gaps in the accumulation of cases due to lack of trial activity in HIV associated cases between 2011 and 2012.
2. I found the VNIa-5 lineage was more widely dispersed in Vietnam than the other 6 lineages.
3. I found there was clustering of 3 lineages: VNIa-4, VNIa-5 and VNIa-32.

4. I found *Cryptococcus neoformans* to be elusive in the environment – I isolated many other cryptococcal species, but only 3 *Cryptococcus neoformans*.
5. Environmental non-*neoformans* cryptococci were more frequently isolated around the spatial cluster of cryptococcal meningitis ST5 cases.
6. ST5 strains from HIV uninfected patients are more virulent than non-ST5 strains in the *G. mellonella* infection model.
7. Clinical isolates (ST4&ST5) are significantly more virulent than environmental isolates (ST4&ST5)
8. The virulence of naive environmental isolates can be up-regulated by a protein/peptide based inter-yeast signaling system derived from virulent isolates.
9. Virulence of such induced isolates is stable over generations.
10. The induced isolate can in turn up-regulate the virulence of the naive isolate.
11. The enhanced virulence of clinical and induced isolates seen in *Galleria* is associated with significantly higher fungal loads in the larvae.
12. 1700 genes were differentially expressed between clinical and environmental isolates. The transcriptional profile of the induced LD2 isolate resembles that of the highly virulent BMD761 strain.
14. Genes annotated to metabolism and cellular biosynthesis were enriched in clinical isolates
15. Besides metabolism and cellular biosynthesis, genes annotated to cell cycle and cell division were enriched in the induced LD2 relative to the naive LD2.
16. Multiple genes associated with virulence factors were up-regulated in clinical and induced environmental isolates.
17. A third of differentially expressed genes between BMD761/Induced LD2 and naive LD2 were hypothetical proteins.

6.3 Discussion

6.3.1 Aim 1: To determine whether there is temporal-spatial clustering of *C. neoformans* by genetic lineage and HIV status in Vietnam

Understanding cryptococcal adaptation to its ecological niche may give insight into how particular lineages can cause disease in immunocompetent humans. As a first

step, I investigated whether there was clustering of cryptococcal meningitis cases. Evidence of disease clusters could provide insight into the ecological habitat of *C. neoformans* in Southern Vietnam where the vast majority of cryptococcal meningitis occurs. I found a spatial cluster spanning Ho Chi Minh city, Dong Nai and Ba Ria-Vung Tau. This clustering is likely driven by HCMC and Ba Ria-Vung Tau, these cities having the highest prevalence of HIV infection in southern Vietnam (1.25% as opposed to 0.5% for the country[183]). HCMC, Dong Nai and Ba Ria-Vung Tau constitute an industrial hub in Vietnam which attracts youths seeking career opportunities. Youths (26-36 years-old) account for 50% of patients infected with cryptococcosis. I do not have the internal migration histories of the cases of cryptococcal meningitis, which, together with the probably prolonged incubation period/latency of cryptococcal infection, confers a complication in interpreting the significance of these clusters. Thus it is possible that the clustering I identified in clinical cases is artefactual. Nonetheless, to my knowledge this is the first study to map and investigate the temporal-spatial clustering of cryptococcal meningitis due to *C. neoformans*, and the first to attempt to determine clustering by genetic lineage. Previously, the ecological niche of the sibling species *C. gattii* causing the outbreak in Vancouver Islands in Canada was modeled. They clustered along the central and southeastern coast of Vancouver Island[170].

I also found spatial clustering of meningitis by infecting lineage. Lineage VNIa-5 infections clustered to the south-west of Ho Chi Minh City. I found clustering of the VNIa-4 lineage to the west alongside the Cambodian border. This is consistent with our understanding of the wider population structure of *C. neoformans* in southeast Asia – this lineage dominates infection in Thailand and Laos[305].

Interestingly, the VNIa-5 lineage had the widest geographical dispersal compared with the other 6 lineages (Figure 2-10). Its expanded range could be explained by its adaptation to multiple different ecosystems in Viet Nam, driven by its phenotypic heterogeneity - Thanh *et al* found VNIa-5 strains to have marked variation within virulence-associated phenotypes [247]. Phenotypic heterogeneity in other yeasts (*S. cerevisae*) facillitates survival in the face of environmental stresses by generating

outliers with exceptionally high fitness[306]. Ecosystems in Viet Nam differ by region at least in terms of tree abundance, elevation, rain fall, temperature and seasonality. There are also significant differences in land use - the Mekong delta is divided by a net of rivers and covered in lush vegetation, mostly rice fields. The neighbouring Southeast is home to industrial tree plantations (rubber, acacia, cassava & cashew). The central highlands, located between 500 and 1500 meters, is famous for cultivation of coffee, tea, white mulberry, and black pepper. The south central coast is a harsh region which suffers storms frequently, and is characterized by the presence of fish farms and culture of non-indigenous animals including sheep, goat and ostrich. However, to determine that the expanded distribution of VNIa-5 that I found through analysis of human cases is real would need environmental sampling. A further option that might provide insight, and be more successful than tree, air and soil sampling, would be analysis of cases derived from veterinary infection. The shorter lifecycles and reduced mobility of animals would negate some of the complications (latency and migration) that interfere with the interpretation of data from humans. However, veterinary medicine is poorly developed in Vietnam and while cryptococcosis in animals is well described, there are currently no epidemiological data from Vietnam.

Interestingly, I did not find clustering of non-HIV associated cryptococcosis cases. This could be explained by the widespread dispersal of VNIa-5 *C. neoformans* isolates (which causes the majority of disease in HIV uninfected patients) or lack of power due to the relatively low number of non-HIV cases available to analyse (59 non-HIV vs 350 HIV). However, since I was able to show clustering of the VNIa-32 lineage with only 12 cases, I think the low power explanation is unlikely, (Figure 2-15), and believe the widespread distribution of VNIa-5 is a more likely explanation. The cluster of HIV-associated cryptococcal meningitis that I identified spanning the whole Mekong River delta and half of Ho Chi Minh city from 2004 to 2006 may be driven by the highest prevalance of people living with HIV in these region in 2005[183].

6.3.2 Aim 2: To identify ecological niches associated with *C. neoformans* lineages

As mentioned above, the movement of people and likely variable latency period of cryptococcal infection are significant complications when trying to understand the ecology and biology of *C. neoformans*. Environmental sampling, which is not complicated by these factors, has the potential to allow a more accurate understanding of spatial distribution, population structure and ecology. I used randomized environmental sampling to ensure that my environmental isolate collection would be representative of the wider population. In particular, I was interested in the relative prevalence of different lineages in relation to their prevalence in clinical isolates, since this ratio could be considered a marker of their relative pathogenicity. Randomized sampling would allow this inference to be drawn robustly. However, I had extremely low rates of isolation of *C. neoformans* using the randomized approach, with recovery rates of 0.06%. Remarkably, I found the recovery rate of non-*neoformans* cryptococci in cluster sites of clinical *C. neoformans* to be five times higher than with random sampling. Morever non-*neoformans* isolates from air were only recovered in the cluster sites of clinical isolates. The high rate of isolation of non-*neoformans* isolates from around cases of human *C. neoformans* disease suggests that, prolonged incubation and latency notwithstanding, there is much in common in the general ecological niches of *Cryptococcus* species. However, the extremely low rate of isolation of pathogenic forms suggests that *Cryptococcus neoformans* may usually be in a state of quiescence in the environment. Such possible quiescent states include spores, desiccated yeasts and the recently discovered novel and quiescent 'viable but non-culturable) forms described by Hommel and colleagues (https://www.biorxiv.org/content/10.1101/552836v1). More intenstive longitudinal sampling, particularly around sites of cases, could be more successful in isolating environmental examples and confirming the ecological niche of *C. neoformans*, as has been done in India and Colombia where an association was found for *C. gattii* with the *Syzygium cumini* and *Terminalia catappa* tree species[172] [26].

6.3.3 Aim 4: To characterize the virulence phenotypes of environmental *Cryptococcus* isolates.

I was interested in potential virulence of environmental isolates, so I characterize their phenotypes using conventional virulence factors including temperature-dependent growth, melanin production, melanin production, capsule size, and ability to survive *ex vivo* CSF. As expected, *C. neoformans* exhibited the most robust expression of such attributes particularly thermotolerance. All *Cryptococcus* can grow at 30^{o}C but only *C. neoformans* can survive human physiological temperature (37^{o}C) and elevated temperature (39^{o}C). This suggested adaptation to human host's physical temperature is the critical virulence attribute for cryptococcal infection. I also challenged *G. mellonella* larvae with these isolates at 30 and 37^{o}C. Surprisngly *Cryptococcus spp* exhibited no significant difference in virulence at 30^{o}C. However *C. neoformans* was significantly more virulent than *C. laurentii* at 37^{o}C given that they grew equally well at this temperature. This implied that thermotolerance is not solely responsible for infection. The finding that only 47% of *C. laurentii* isolates in my study can grow at 37^{o}C agree with rare report of *C. laurentii* infection.

C. laurentii rarely cause opportunistic disease involving skin, lung, bloodstream and the central nervous system[211]. Infection in humans caused by other cryptococcal species (*C. heveanensis*, *C. rajasthanensis* and unidentified *Cryptococcus*) has not been reported.

My phenotypic assays suggest that larva infection model with exposure at 37^{o}C could be used to investigate potential pathogenicity of environmental fungal pathogens instead of relying on individual virulence factors.

6.3.4 Aim 5: To compare the virulence between and within different genotypes using the *G. mellonella* model.

Two decades ago, *G. mellonella* emerged as an alternative infection model to study microbial pathogenesis. The insect immune response shares functional and structural homology to the innate immune response of vertebrates, with insect hemolymph containing at least six types of hemocytes which function similarly to mammalian phagocytes[229]. *G. mellonella* has been shown to provide comparable

data to murine models of infection for a number of microbial species including *Pseudomonas aeruginosa* [163], *Bacillus cereus* and *B. thuringiensis*[229]. With fungi, the *G. mellonella* infection model can distinguish pathogenic and non-pathogenic *Candida albicans* isolates and there is a positive correlation between the pathogenicity of *C. albicans* in mice and larvae. Subsequently, an association between toxin production and the virulence of *Aspergillus fumigatus* was established in the model[165].

Day *et al* previously found that *C. neoformans* ST5 isolates cause cryptococcal infection in immunocompetent individuals more frequently than non-ST5 isolates[86]. The lineage ST5 is of great interest because it predominates in China and neighbouring countries in north eastern Asia (Japan, Korea) where disease in HIV-uninfected patients has been frequently reported. Given this host distribution, I hypothesized that ST5 strains have increased pathogenicity – ability to cause disease - compared with non-ST5 strains. Using the *Galleria* infection model, I found ST5 isolates derived from HIV-uninfected patients were significantly more virulent than ST4 isolates derived from HIV-patients. Moreover, this model can differentiate naive environmental isolates from virulent clinical isolates (ST4 and ST5). The difference in virulence between clinical and environmental *C. neoformans* isolates (serotype A) has previously been demonstrated in the mouse infection model[62]. The most exciting part of my work was the finding of an 'ecological imprint' on strains. Strains derived from HIV uninfected apparently immunocompetent patients are more virulent than those from HIV infected (immunosuppressed) patients, which are in turn more virulent than strains derived from the environment. This virulence phenotype is extremely stable and fixed over generations of infection and purification. Moreover, while increased virulence is not 'learnt' by repeated infection in the model, it can be 'taught' by hypervirulent, lineage 5, strains to less virulent strains. Once upregulated, that virulence phenotype is again fixed over generations, and the upregulated strain itself can now upregulate the virulence of other strains – i.e the effect is transmissible. My experiments suggest that the phenomenon is driven by secreted peptides/proteins. The phenomenon offers a fascinating avenue of research to take forward, with the potential identification of novel drug targets. Interrupting inter-

yeast communication that is at the heart of regulation of virulence is an exciting prospect for the control of infection.

The whole genome sequence-derived phylogeny of *C. neoformans* generated within our research group at OUCRU (Appendix B) shows that VNIa-5 strains from immuncompetent patients, that have the hypervirulent phenotype, are dispersed throughout the phylogeny of that lineage. This suggests that all VNIa-5 isolates have the potential to cause disease in immunocompetent hosts.

Differences in the virulence of clinical isolates versus environmental isolates of strains of the same molecular type has been described before in the mouse infection model[62]. Litvintseva *et al* found the vast majority of environmental strains to be avirulent, whereas 70% of clinical strains were virulent. Together with my data, this begs the big question regarding how do naive environmental isolates become pathogenic in the environment, in order to cause human infection. Do they evolve to become pathogenic through their interaction with natural predators (amoeba, nematode, etc.) or through incubation in human (or other mammalian) hosts? I proved that passage of naive *C. neoformans* through six generations of larvae did not enhance virulence of passaged isolates, suggesting at least that it is not the innate immune system that drives virulence up-regulation. However, in my experiments, yeasts extracted from larval hemolymph were cultured on Sabouraud agar before infecting the next batch of larvae. It could be argued that this 'interruption', outside the immune environment, may result in a reversion of virulence. An alternative approach that would mitigate this potential effect would be to perform *in vivo* passage. In this case, yeasts extracted from larval hemolymph are injected directly into the next batch of larvae without prior culture, so they would in theory maintain virulence acquired during interaction with hemocytes. However, the fact that the upregulated virulence following growth in culture filtrate from a hypervirulent strain is also fixed despite growth on agar plates between larvae infections, and even following freezing, suggests that this 'interruption-reversion' hypothesis is wrong. In fact, the durability of virulence upregulation is intriguing, suggesting epigenetic effect such as chromatin remodeling, and is an additional opportunity for research in the future, alongside elucidating the mechanics of inter-yeast signaling itself.

A further explanation for discrepancy in virulence between clinical and environmental isolates is that pathogenicity/virulence is primarily a quality of spores rather than yeast cells, and a 'pathogenicity' switch is thrown when a spore ends in a host-like environment. An obvious trigger for this could be thermotolerance in mammals. It would be interesting to induce sporulation in clinical and environmental isolates and see whether there is variability in the virulence in infection models where the spore is used as the inoculum.

6.3.5 Aims 6 and 7: To define transcriptional pathways and to identify genes associated with variability in virulence between isolates

I found infection with clinical isolates and induced environmental isolates led to moderately higher fungal burdens in the *Galleria* model compared with infections due to the environmental strain. Whether this higher fungal burden is an expression of higher virulence, or the cause of increased virulence (presumably as a function of an increased rate of growth), is not clear. Higher fungal burdens could be due to greater rates of proliferation or could reflect better survival/evasion of the host immune response in the *Galleria* compared with naive isolates. However, there was no change in *in vitro* growth rates of the environmental strain following induction (appendix C). The higher fungal burden of the BMD761 and induced *C. neoformans* isolates seen in larvae was associated with upregulation of virulence-associated genes (virulence factors, kinases, transporters, transcription factors) and enrichment of genes annotated to metabolism, cellular biosynthesis, and cell cycle, consistent with higher growth rates. Around 260 hypothetical proteins were differentially regulated in induced LD2 and naive LD2. Presumably, somewhere within this set of genes is a single gene or group of genes that function as intercellular signals to upregulate virulence. It is possible that some of these upregulated genes I observed are artefactual – I sequenced RNA from *C. neoformans* cells which had been revived in YPD rather than *C. neoformans* cells extracted directly from larvae. However, virulence induction did appear to be an extremely robust *in vitro* phenomenon. RNAseq on yeast cells derived directly from hemolymph might provide cleaner transcriptional data, but separating yeast cells without contamination of hemocyte debris is challenging. It is an experiment I would consider taking forward in the

future. However, a more fruitful approach is likely to be a proteomic approach, particularly since the relationship between protein/peptide expression and trsncription is not necessarily linear. Following the identification of protein/peptide candidates, the advent of CRISPR technology, reliably deliverable for *C. neoformans* using electroporation, delivers the tools to allow further interrogation of my isolates and gene candidates.

6.4 Future directions

If I were to take my work on *Cryptococcus* further forward, I would be most interested in better understanding the mechanism of inter-yeast signaling, and how this upregulates pathogenesis.

My works suggest that inter-yeast signaling is certainly in someway mediated through secreted proteins, and as discussed in chapter 4 recent work suggests that extracellular vesicles may facilitate inter-yeast communication in *C. neoformans* and may have a role in the regulation of virulence of the related species *C. gattii*. I would be interested in determining whether vesicles are mediating the effect I have described in my isolates, which I believe is plausible based on the following evidence. First, the constituency of vesicles is complex. They contain cell membrane, proteins, which may be either within the membrane or cross the membrane, and they may also contain small RNAs. Rodrigues *et al* reported *C. neoformans* extracellular vesicles containing plentiful protein components associated with virulence[119]. Second, Bielska *et al* found *C. gattii* extracellular vesicles can mediate intercellular communication between outbreak and non-outbreak strains, resulting in probable up-regulation of virulence. They found that the intracellular proliferation rate (IPR) of non-outbreak strains in murine macrophages was significantly enhanced when these strains was co-cultured with outbreak strains[230]. They went on to show that this is regulated by extracellular vesicles (EV) derived from the outbreak strain. They believed that the mechanism was mediated through the uptake of EVs by the host macrophage, with transport to the fungal phagosome where they triggered proliferation of non-outbreak cells. The two EV components essential for inducing virulence of non-outbreak strains were reported as protein and RNA[230].

My work differs in that it shows that macrophages are not essential to the inter-yeast communication – I demonstrate up-regulation occurring in *in vitro* culture. However, I have not demonstrated the presence of extracellular vesicles, but could gather some supporting information for this with some fairly simple experiments.

Wolf *et al* found albumin disrupts bacterial and fungal vesicles[307]. Therefore, as a first step, I would destabilize EV in culture filtrate by adding albumin to the culture filtrate and test whether this disrupts the induction of virulence in the *Galleria* model.

A second experiment would be to use ultracentrifugation to separate the potential vesicle containing fraction from the culture filtrate (CF). Then I would use the (EV-containing) pellet and (EV-free) supernatant and determine which of these has the ability to induce virulence in naive LD2, again testing the phenotype using the *Galleria* model. Finally, I could determine the presence of EVs in culture filtrate (and in ultracentrifuged CF) using transmission electron microscopy. However, I do not have this last facility in Vietnam.

Having demonstrated that vesicles are implicated in the phenomenon, I would then wish to purify and fractionate them, and use mass spectrometry to define the proteins present. If I could fractionate vesicles by size, which can be done with filtration columns, then I would be able to test the effect of individual fractions on induction.

Of note, when I treated culture filtrate with nucleases, there was no effect on its ability to upregulate virulence. EVs can contain small RNAs, and even DNA. I did not find any evidence that nucleic acids mediate the induction effect of culture filtrate. However, it may be that small RNAs, contained within vesicles, are somewhat protected against the action of nucleases. Therefore, having purified EVs (assuming they underly the phenomenon) I would be interested in small RNA-seq to determine whether RNA is present and what form it takes.

In Chapter 5, I found a large number of virulence-associated genes that were upregulated when the environmental isolate was induced to become more virulent. I have discussed the biological plausibility of this differential expression in the discussion of chapter 5. Differentially expressed genes between clinical/induced

isolates and environmental isolates could also be validated by quantification of their expression using real-time reverse-transcription PCR.

To test whether my inferences are correct, I would need to set up knock-out experiments. The advent of CRISPR-Cas9 system for *Cryptococcus* enables the precise development of relevant knockouts[308]. The challenge would be identifying which genes to knockout first. Here, because it is increasingly clear that the relationship between gene expression and phenotype is not simply linear, I would be guided by the results of the proteomic analyses of the culture filtrate/EVs as described above, in choosing which genes to knock-out. Having made knockouts of the BMD761 hypervirulent isolate, I would then use the *Galleria* model to evaluate the effect of mutant BMD761 culture filtrate on the naive LD2 isolate. Further experiments to define the exact mechanism through which increased virulence is mediated would have to be guided by the genes implicated from the knockout experiments. If, for example, these were genes involved in sugar transport (as described in chapter 5), then simple experiments looking at growth efficiency in differing relevant sugar concentrations might be informative.

6.5 Conclusion

In summary, I found some evidence of clustering of *Cryptococcus* species within the wider environment, but was unable to identify specific niches associated with particular lineages because of poor recovery rates. The increased isolation rates of any *Cryptococcus* species from around clinical cases indicates that the ecological niches are highly similar and likely differ at the micro- rather than macro- level.

The identification of a phenomenon suggestive of inter-yeast signaling that leads to upregulation of virulence is an exciting therapeutic prospect and something I would be interested to take forward in the future.

Appendix A

List of environmental isolates in Vietnam

Isolate ID	Source	Type of sample	Place of sampling	Method	Species
NT466	Delonix regia	swab	Nhon Trach	random	C. heveanensis
15438	Hopea odorata	swab	HCMC	random	C. heveanensis
VAT550	Khaya senegalensis	swab	HCMC	random	C. rajasthanensis
NT4419	Mimosa pigra	swab	Nhon Trach	random	Undetermined cryptococcus
NT6628	Cassia fistula	swab	Nhon Trach	random	C. heveanensis
NT11127	Tamarindus indica	swab	Nhon Trach	random	C. heveanensis
NT6620	Lagerstroemia speciosa	swab	Nhon Trach	random	C. heveanensis
17479	unidentified	swab	HCMC	random	C. heveanensis
GD299	Khaya senegalensis	swab	HCMC	random	C. laurentii
6776	Tamarindus indica	rotten wood	HCMC	random	C. laurentii
599	Dipterocarpus alatus	swab	HCMC	random	C. heveanensis
4367	Terminalia catappa	rotten wood	HCMC	random	C. heveanensis
25663	Lagerstroemia speciosa	soil	HCMC	random	C. laurentii
2563	Alstonia scholaris	swab	HCMC	random	C. heveanensis
3602	Delonix regia	swab	HCMC	random	C. heveanensis
NT3039	Peltophorum pterocarpum	swab	Nhon Trach	random	C. laurentii
NT1213	Cassia fistula	swab	Nhon Trach	random	C. heveanensis
NT11534	Hevea brasiliensis	swab	Nhon Trach	random	C. laurentii
GeckoJD	Guano	gecko guano	HCMC	target	C. heveanensis
AirconJD	air conditioner filter	swab	HCMC	target	C. rajasthanensis
WallJD	wall	swab	HCMC	target	C. laurentii

NT7533	Acacia auriculiformis	swab	Nhon Trach	random	C. neoformans
6338	Dipterocarpus alatus	swab	HCMC	random	C. laurentii
6350	Delonix regia	swab	HCMC	random	C. heveanensis
2697	Samanea saman	swab	HCMC	random	C. heveanensis
11336	Delonix regia	swab	HCMC	random	C. heveanensis
11516	Artocarpus altilis	swab	HCMC	random	C. heveanensis
SS153	Mimusops elengi	swab	HCMC	target	C. laurentii
SS41	Artocarpus heterophyllus	air	HCMC	target	C. laurentii
SS37	Peltophorum pterocarpum	air	HCMC	target	C. heveanensis
SS67	Ficus rumphii	swab	HCMC	target	C. heveanensis
SS97	Delonix regia	swab	HCMC	target	C. heveanensis
SS105	unidentified	swab	HCMC	target	C. heveanensis
SS118	Rhizophora apiculata	swab	HCMC	target	C. heveanensis
SS77	Bougainvillea spectabilis	swab	HCMC	target	C. laurentii
SS54	Pterocarpus macrocarpus	swab	HCMC	target	C. heveanensis
NT4268	Cassia fistula	swab	Nhon Trach	random	C. heveanensis
NT3960	tree stump	swab	Nhon Trach	random	C. laurentii
NT3176	tree stump	swab	Nhon Trach	random	C. laurentii
LTR82	Hopea odorata	swab	HCMC	random	C. heveanensis
LTR73A	Cassia fistula	swab	HCMC	random	C. heveanensis
LTR73B	Cassia fistula	swab	HCMC	random	C. rajasthanensis
LD1	Hopea odorata	swab	HCMC	random	C. neoformans
LD2	Hopea odorata	swab	HCMC	random	C. neoformans
7908	Hopea odorata	swab	HCMC	random	C. heveanensis
5309	Khaya senegalensis	swab	HCMC	random	C. heveanensis
5521	Hopea odorata	swab	HCMC	random	C. heveanensis
24326	Peltophorum pterocarpum	swab	HCMC	random	C. heveanensis
4918	Dipterocarpus alatus	swab	HCMC	random	C. heveanensis
1295	unidentified	soil	HCMC	random	C. heveanensis
10127	Dipterocarpus	swab	HCMC	random	C. heveanensis

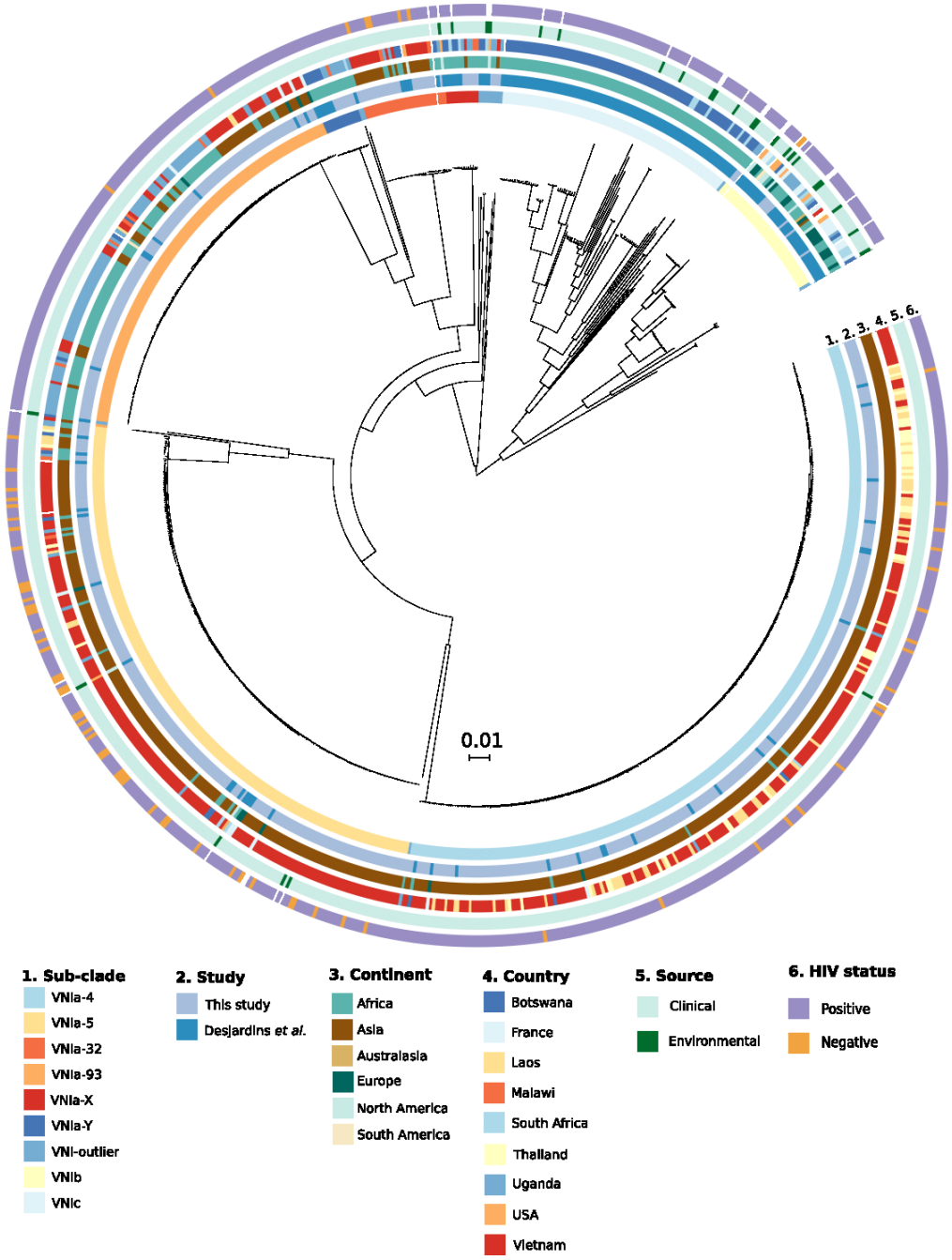
	alatus				
9940	Cinnamomum camphora	swab	HCMC	random	C. heveanensis
12896	Azadirachta indica	swab	HCMC	random	C. heveanensis
14300	Mimusops elengi	swab	HCMC	random	C. heveanensis
11888	Delonix regia	swab	HCMC	random	C. heveanensis
4917	Samanea saman	swab	HCMC	random	C. heveanensis
10126	Dipterocarpus alatus	swab	HCMC	random	C. heveanensis
14482	Delonix regia	swab	HCMC	random	C. laurentii
15077	Delonix regia	swab	HCMC	random	C. heveanensis
19104	Delonix regia	swab	HCMC	random	C. heveanensis
TD135	unidentified	swab	HCMC	random	C. heveanensis
6865	Delonix regia	swab	HCMC	random	C. heveanensis
9658	Delonix regia	swab	HCMC	random	C. heveanensis
5063	Acacia mangium	swab	HCMC	random	C. heveanensis
1021	Cassia fistula	swab	HCMC	random	C. heveanensis
9494	Peltophorum pterocarpum	swab	HCMC	random	C. heveanensis
5052	Samanea saman	swab	HCMC	random	C. heveanensis
853	Peltophorum pterocarpum	swab	HCMC	random	C. heveanensis
661	Cassia fistula	swab	HCMC	random	C. heveanensis
1266	Mimusops elengi	swab	HCMC	random	C. heveanensis
1419	Cassia fistula	swab	HCMC	random	C. heveanensis
3657	Samanea saman	swab	HCMC	random	C. heveanensis
4827	Peltophorum pterocarpum	soil	HCMC	random	C. heveanensis
13118	Delonix regia	swab	HCMC	random	C. heveanensis
39469	Hopea odorata	swab	HCMC	random	C. heveanensis
258	Eucalyptus globulus	swab	HCMC	random	C. heveanensis
15091	Delonix regia	swab	HCMC	random	C. heveanensis
11873	unidentified	swab	HCMC	random	C. heveanensis
1810	Pterocarpus macrocarpus	swab	HCMC	random	C. heveanensis
6252	Cassia fistula	swab	HCMC	random	C. heveanensis
23478	Muntingia calabura	swab	HCMC	random	C. heveanensis
25498	Lagerstroemia	soil	HCMC	random	C. heveanensis

	speciosa				
27498	Peltophorum pterocarpum	swab	HCMC	random	C. heveanensis
37969	Muntingia calabura	swab	HCMC	random	C. laurentii
40147	Ravenala madagascariensis	swab	HCMC	random	C. heveanensis
NT2975	Manihot esculenta	swab	Nhon Trach	random	Undetermined cryptococcus
1077	Pterocarpus macrocarpus	swab	HCMC	random	C. heveanensis
1474	Samanea saman	swab	HCMC	random	C. heveanensis
5092	Peltophorum pterocarpum	swab	HCMC	random	C. heveanensis
5486	Hopea odorata	swab	HCMC	random	C. heveanensis
6285	Cassia fistula	swab	HCMC	random	C. heveanensis
8870	Delonix regia	soil	HCMC	random	C. heveanensis
6495	Hopea odorata	swab	HCMC	random	C. laurentii
38262	Acacia auriculiformis	swab	HCMC	random	C. heveanensis
37796	Delonix regia	swab	HCMC	random	C. heveanensis
1673	Samanea saman	swab	HCMC	random	C. rajasthanensis
35178	Delonix regia	swab	HCMC	random	C. heveanensis
32766	Delonix regia	swab	HCMC	random	C. heveanensis
VAT	tree stump	swab	HCMC	random	C. laurentii
SS160a	Lagerstroemia speciosa	swab	Vung Tau	target	C. heveanensis
SS160b	Plumeria rubra	swab	Vung Tau	target	C. heveanensis
31049	Khaya senegalensis	swab	HCMC	random	C. laurentii
27450	Lagerstroemia speciosa	swab	HCMC	random	C. heveanensis
34467	Muntingia calabura	swab	HCMC	random	C. heveanensis
39328	Tamarindus indica	swab	HCMC	random	C. laurentii
38296	Peltophorum pterocarpum	soil	HCMC	random	C. heveanensis
14797	Mimusops elengi	swab	HCMC	random	C. laurentii
4067	Dipterocarpus alatus	swab	HCMC	target	C. heveanensis
10771	Pterocarpus macrocarpus	swab	HCMC	random	C. heveanensis

NT4243	Peltophorum pterocarpum	swab	Nhon Trach	random	C. heveanensis
TXS377	Peltophorum pterocarpum	swab	HCMC	random	C. heveanensis
SS74	Hopea odorata	swab	Tien Giang	target	C. rajasthanensis
35244	Peltophorum pterocarpum	soil	HCMC	random	C. heveanensis
36053	Peltophorum pterocarpum	swab	HCMC	random	C. heveanensis
33388	Tamarindus indica	swab	HCMC	random	C. heveanensis
26977	Delonix regia	swab	HCMC	random	C. heveanensis
25340	Cassia fistula	swab	HCMC	random	C. heveanensis
23747	Delonix regia	swab	HCMC	random	C. heveanensis
18740	Delonix regia	swab	HCMC	random	C. heveanensis
NT6004	Lagerstroemia speciosa	swab	Nhon Trach	random	C. rajasthanensis
SG18680	Melia azedarach	swab	HCMC	random	C. heveanensis
35386	unidentified	swab	HCMC	random	C. heveanensis
6289	unidentified	swab	HCMC	random	C. heveanensis

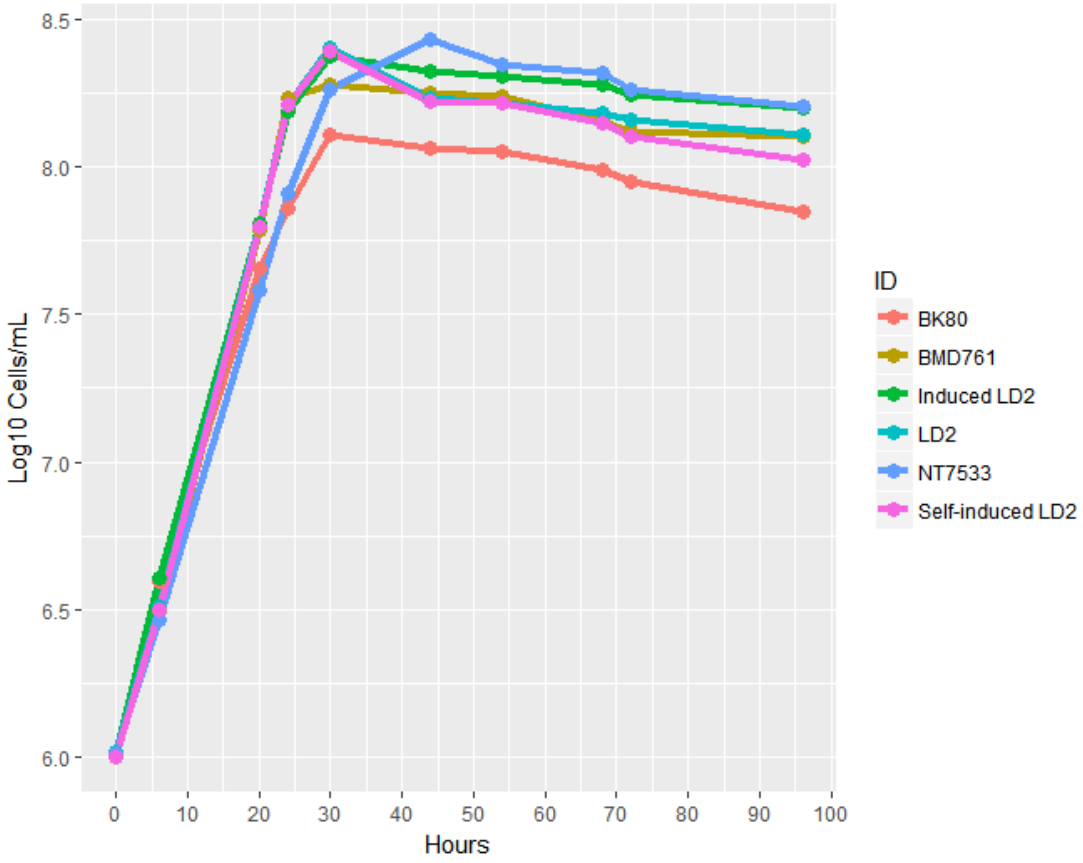
Appendix B

A whole genome SNP phylogeny of 679 *C. neoformans* VNI isolates (https://www.biorxiv.org/content/biorxiv/early/2018/06/28/356816.full.pdf)



Appendix C

Growth curves of strains used in the chapter 5



In vitro growth rates (from 0-20 hours) at 30^{o}C were similar between isolates (P=0.26, Kruskal-Wallis rank sum test).

Appendix D

Protocol for *G. mellonella* infection model

1. Purpose

This protocol describes the usage of *Galleria mellonella* larvae as infection model to study pathogenicity of *Cryptococcus* species.

2. Principle

Cryptococcus species vary in their pathogenicity between species and even between different genotypes. Insect larvae can play as a host to study pathogenicity of Cryptococcus pathogen. Larvae will be infected with the same inoculum of *Cryptococcus* species/genotypes of interest. All infected larvae are incubated for ten days at the same temperature. Larvae survival is observed daily using death as an endpoint. Survival curves plotted after 10 days could tell if the pathogens of interest significantly differ in their pathogenicity.

3. Scope and responsibilities

This procedure will be carried out within the BSL2 cabinet of mycology lab, OUCRU, HCMC. Only trained staff members of the CNS/HIV group are allowed to perform the procedure.

4. Reference

Mylonakis E, Moreno R, El JB, Idnurm A, Heitman J, Stephen B, et al. Galleria mellonella as a model system to study Cryptococcus neoformans pathogenesis. Infect Immun. 2005;73(7):3842–2850.

Website: https://www.jove.com/video/50964/use-galleria-mellonella-as-model-organism-to-study-legionella

5. Specimens

- Galleria mellonella larvae (250-300 mg)
- *Cryptococcus* species

6. Materials/ Equipments used

- 1x PBS (phosphate-buffered saline)
- Light microscope fitted with 40x objective
- Forceps
- Glass slide & slide cover
- 10 micrometer Hamilton syringe
- Petri dish
- Rubber and leather gloves
- YPD broth (1% yeast extract, 2%peptone and 2% dextrose)
- Sabouraud and BHI (brain heart infusion) agar plates
- 0.2 micrometer pore size filter
- 10 mL syringes
- 10 micromiter loop
- 70 % alcohol
- Bleach (chlorine-based)
- Cellometer (Nexcelom)
- 1.5 mL microtube
- Centrifuge

7. Procedures

Prepare larvae

1) Larvae, kept in 16 ^{o}C cooler, are taken out and left at room temperature for at least 2 hours before infection.
2) Larvae are dispensed into petri dishes (20 larva/dish) prior to injection
3) Discard any larvae with grey markings (indicator for ill larvae)

Fungal inoculum preparation

4) Fungi are cultured in YPD broth between 24 to 48 hours in shaking incubator (150-200 rpm) at 30 ^{o}C
5) Spin the culture at 8000 rpm for 1 minute
6) Wash the pellet with PBS and then resuspend pellet in PBS
7) Quantitate fungal cells using the cellometer.
8) Adjust a density of 10^8 cells/mL, then 10 µL of inoculum (10^6 cells/larva) is ready for injection

Sterility test

Sterility of culture and PBS can be checked by inoculating 10 µL of culture into Sabouraud and BHI. Incubate plates at 30 ^{o}C for 3 days.

Infection

9) Clean Hamilton syringes using 10% bleach by drawing and pushing out for 3 times followed by 100% ethanol and 1X PBS.
10) The needle and syringe are cleaned between fungal strains
11) Use a cotton bud soaked in 70 % ethanol to swab the larva pro-legs prior to injection.
12) Inject 10 µL of inoculum (1.0x106 cells per larvae) into the last left pro-leg. If there is tension then adjust the position slightly. It should enter the larvae easily. Don't push hard and fast, needle is sharp and easy to go through the larvae body.
13) After injection, put larva into a petri dish covered with tissue for a few minute to stop hemolymph leakage
14) Also inject 10 µL PBS as control
15) Transfer infected larvae into a new petri dish (maximum 20 larvae/plate). Use tape to fasten both sides of petri dishes to prevent larvae from escaping.
16) Put petri dishes in perforated platic boxes or plastic box covered with aluminum foil and incubate at 37 ^{o}C incubator for 10 days
17) NOTE: a leather glove must be worn during injection

Observation after injection

18) Larvae are checked daily for survival by physical stimuli with forceps. Remove cocoons if the larvae can't be observed. Normally they start make a little silk immediately after injection.
19) Dead larvae are removed from petri dish. Put them in ziplock bags and discard in a biological harzard bin. The number of dead larvae is recorded.
20) Analyse result by plotting survival curves

8. Quality control

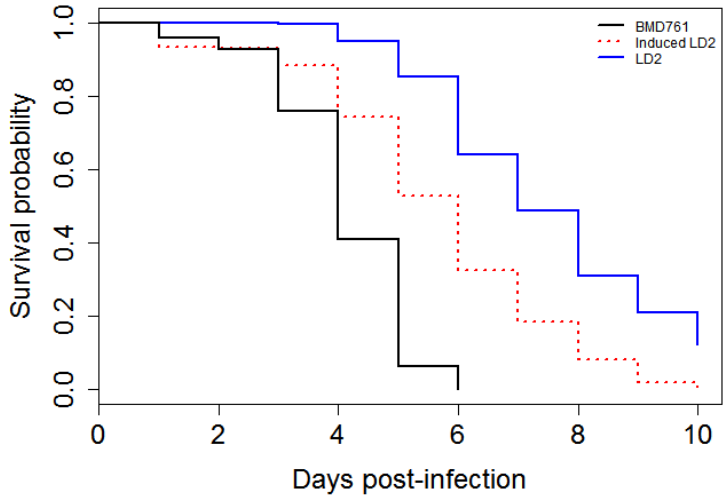
- Quantitate inoculum after infection by inoculating onto Sabouraud agar
- All petri dishes are encoded by another labmate to avoid bias

9. Safety

- PPE should be worn at all times while working (a minimum of lab coats, closed shoes, gloves)
- Place any waste from the procedure in the appropriate hazardous waste bins.

Appendix E

Survival curves derived from pooled data from 6 independent *Galleria mellonella* larva infection experiments. Larvae were infected with one of 3 *Cryptococcus neoformans* isolates: BMD 761, naive LD2 or induced LD2. Survival data were pooled by infecting isolate. BMD761 is an isolate derived from an HIV uninfected and apparently immunocompetent human patient; LD2 is an environmental isolate of the same genetic lineage as BMD 761 (VNIa-5). Induced LD2 is the LD2 isolate following its culture in broth augmented with sterile culture filtrate from BMD761 culture. The sample size is 240 larvae per arm. Infection with the induced LD2 isolate resulted in a 2.5 fold increase in the hazard of death (i.e. Hazard Ratio = 2.5) compared with infection with naïve LD2 (Cox proportional hazards model, P <0.0001).



Bibliography

1. Rajasingham, R. *et al.* Global Burden of Disease of HIV-Associated Cryptococcal Meningitis: an Updated Analysis. *Lancet Infect. Dis.* **3099,** 1–9 (2017).
2. Feldmesser, M., Tucker, S. & Casadevall, A. Intracellular parasitism of Cryptococcus neoformans. *Trends Microbiol.* **9,** 273–278 (2001).
3. Steenbergen, J. N., Shuman, H. A. & Casadevall, A. Cryptococcus neoformans interactions with amoebae suggest an explanation for its virulence and intracellular pathogenic strategy in macrophages. *Proc. Natl. Acad. Sci. U. S. A.* **98,** 15245–15250 (2001).
4. Mylonakis, E., Ausubel, F. M., Perfect, J. R., Heitman, J. & Calderwood, S. B. Killing of Caenorhabditis elegans by Cryptococcus neoformans as a model of yeast pathogenesis. *Proc. Natl. Acad. Sci.* **99,** 15675–15680 (2002).
5. Steenbergen, J. N., Nosanchuk, J. D., Stephanie, D., Casadevall, A. & Malliaris, S. D. Cryptococcus neoformans Virulence Is Enhanced after Growth in the Genetically Malleable Host Dictyostelium discoideum. *Infect. Immun.* **71,** 4862–4872 (2003).
6. Chaturvedi, V. & Chaturvedi, S. Cryptococcus gattii : a resurgent fungal pathogen. *Trends Microbiol.* **19,** (2011).
7. *Cryptococcus: from human pathogen to model yeast*. *ASM Press* (2011). doi:10.1016/S1473-3099(11)70140-2
8. Franzot, S. P., Salkin, I. F. & Casadevall, a. Cryptococcus neoformans var. grubii: separate varietal status for Cryptococcus neoformans serotype A isolates. *J. Clin. Microbiol.* **37,** 838–40 (1999).
9. Meyer, W., Castañeda, A., Jackson, S., Huynh, M. & Castañeda, E. Molecular typing of IberoAmerican Cryptococcus neoformans isolates. *Emerg. Infect. Dis.* **9,** 189–95 (2003).
10. Bartlett, K. H. *et al.* A Decade of Experience : Cryptococcus gattii in British Columbia. *Mycopathologia* **173,** 311–319 (2012).
11. Trilles, L., Meyer, W., Wanke, B., Guarro, J. & Ra, M. L. A. Z. É. Correlation of antifungal susceptibility and molecular type within the Cryptococcus

neoformans / C . gattii species complex. *Med. Mycol.* **50,** 328–332 (2012).

12. Khan, Z. U., Randhawa, H. S., Kowshik, T., Chowdhary, A. & Chandy, R. Antifungal susceptibility of Cryptococcus neoformans and Cryptococcus gattii isolates from decayed wood of trunk hollows of Ficus religiosa and Syzygium cumini trees in north-western India. *J. Antimicrob. Chemother.* **60,** 312–6 (2007).
13. Meyer, W. *et al.* Consensus multi-locus sequence typing scheme for Cryptococcus neoformans and Cryptococcus gattii. *Med. Mycol.* **47,** 561–70 (2009).
14. Byrnes, E. J. *et al.* Emergence and pathogenicity of highly virulent Cryptococcus gattii genotypes in the northwest United States. *PLoS Pathog.* **6,** (2010).
15. Ma, H. & May, R. C. *Virulence in Cryptococcus Species*. *Advances in Applied Microbiology* **67,** (2009).
16. Desjardins, C. *et al.* Population Genomics And The Evolution of Virulence In The Fungal Pathogen Cryptococcus neoformans. *Genome Res.* **27,** 1–13 (2017).
17. Lin, X., Hull, C. & Heitman, J. Sexual reproduction between partners of the same mating type in Cryptococcus neoformans. *Nature* **434,** 1017–1021 (2005).
18. Velagapudi, R., Hsueh, Y.-P., Geunes-Boyer, S., Wright, J. R. & Heitman, J. Spores as infectious propagules of Cryptococcus neoformans. *Infect. Immun.* **77,** 4345–55 (2009).
19. Botts, M. R., Giles, S. S., Gates, M. A., Kozel, T. R. & Hull, C. M. Isolation and Characterization of Cryptococcus neoformans Spores Reveal a Critical Role for Capsule Biosynthesis Genes in Spore Biogenesis. *Eukaryot Cell* **8,** 595–605 (2009).
20. Lin, X. & Heitman, J. The biology of the Cryptococcus neoformans species complex. *Annu. Rev. Microbiol.* **60,** 69–105 (2006).
21. Bovers, M. *et al.* Unique hybrids between the fungal pathogens Cryptococcus neoformans and Cryptococcus gattii. *FEMS Yeast Res.* **6,** 599–607 (2006).

22. Bovers, M. *et al.* AIDS patient death caused by novel Cryptococcus neoformans x C. gattii hybrid. *Emerg. Infect. Dis.* **14,** 1105–8 (2008).
23. Lin, X. Cryptococcus neoformans: morphogenesis, infection, and evolution. *Infect. Genet. Evol.* **9,** 401–16 (2009).
24. Litvintseva, A. P., Lin, X., Templeton, I., Heitman, J. & Mitchell, T. G. Many globally isolated AD hybrid strains of Cryptococcus neoformans originated in Africa. *PLoS Pathog.* **3,** (2007).
25. Nielsen, K., De Obaldia, A. L. & Heitman, J. Cryptococcus neoformans mates on pigeon guano: implications for the realized ecological niche and globalization. *Eukaryot. Cell* **6,** 949–59 (2007).
26. Chowdhary, A., Rhandhawa, H. S., Prakash, A. & Meis, J. F. Environmental prevalence of Cryptococcus neoformans and Cryptococcus gattii in India: an update. *Crit. Rev. Microbiol.* **38,** 1–16 (2012).
27. Cogliati, M. Global Molecular Epidemiology of Cryptococcus neoformans and Cryptococcus gattii : An Atlas of the Molecular Types. *Scientifica (Cairo).* **2013,** (2013).
28. Litvintseva, A. P., Thakur, R., Vilgalys, R. & Mitchell, T. G. Multilocus sequence typing reveals three genetic subpopulations of Cryptococcus neoformans var. grubii (serotype A), including a unique population in Botswana. *Genetics* **172,** 2223–38 (2006).
29. Beale, M. A. *et al.* Genotypic Diversity Is Associated with Clinical Outcome and Phenotype in Cryptococcal Meningitis across Southern Africa. *PLoS Negl. Trop. Dis.* **9,** 1–18 (2015).
30. Martinez, L. R. & Garcia-rivera, J. Cryptococcus neoformans var . neoformans (Serotype D) Strains Are More Susceptible to Heat than C . neoformans var . grubii (Serotype A) Strains. *J. Clin. Microbiol.* **39,** 3365–3367 (2001).
31. May, R. C., Stone, N. R. H., Wiesner, D. L., Bicanic, T. & Nielsen, K. Cryptococcus : from environmental saprophyte to global pathogen. *Nat. Rev. Microbiol.* (2015). doi:10.1038/nrmicro.2015.6
32. Coelho, C., Bocca, A. L. & Casadevall, A. *The Tools for Virulence of Cryptococcus neoformans*. *Advances in Applied Microbiology* **87,** (2014).

33. Derengowski, L. S., Paes, C., Albuquerque, P., Tavares, H. F. P. & Fernandes, L. The Transcriptional Response of Cryptococcus neoformans to Ingestion by Acanthamoeba castellanii and Macrophages Provides Insights into the evolutionary adaptation to the mamalian host. *Eukaryot. Cell* **12,** 761–774 (2013).
34. Jung, K.-W. *et al.* Systematic functional profiling of transcription factor networks in Cryptococcus neoformans. *Nat. Commun.* **6,** 6757 (2015).
35. Lee, K.-T. *et al.* Systematic functional analysis of kinases in the fungal pathogen Cryptococcus neoformans. *Nat. Commun.* **7,** 12766 (2016).
36. Magditch, D. A., Liu, T., Xue, C. & Idnurm, A. DNA Mutations Mediate Microevolution between Host- Adapted Forms of the Pathogenic Fungus Cryptococcus neoformans. *PLOS Pathog.* **8,** (2012).
37. Franzot, S. P. *et al.* Microevolution of a Standard Strain of Cryptococcus neoformans Resulting in Differences in Virulence and Other Phenotypes. **66,** 89–97 (1998).
38. Chen, Y. *et al.* Microevolution of Serial Clinical Isolates of C. neoformans var grubii and C. gattii. **8,** 1–18 (2017).
39. Giles, S. S. *et al.* Elucidating the Pathogenesis of Spores from the Human Fungal Pathogen Cryptococcus neoformans. *Infect. Immun.* **77,** 3491–3500 (2009).
40. Datta, K. *et al.* Spread of Cryptococcus gattii into Pacific Northwest region of the United States. *Emerg. Infect. Dis.* **15,** 1185–91 (2009).
41. Kidd, S. E. *et al.* A rare genotype of Cryptococcus gattii caused the cryptococcosis outbreak on Vancouver Island (British Columbia, Canada). *Proc. Natl. Acad. Sci. U. S. A.* **101,** 17258–63 (2004).
42. Ma, H. *et al.* The fatal fungal outbreak on Vancouver Island is characterized by enhanced intracellular parasitism driven by mitochondrial regulation. *Proc. Natl. Acad. Sci.* **106,** (2009).
43. Fraser, J. *et al.* Same-sex mating and the origin of the Vancouver Island Cryptococcus gattii outbreak. *Nature* **437,** 1360–4 (2005).
44. Hagen, F. *et al.* Ancient Dispersal of the Human Fungal Pathogen Cryptococcus

gattii from the Amazon Rainforest. *PLoS One* **8,** (2013).

45. Hagen, F. *et al.* Recognition of seven species in the *Cryptococcus gattii*/*Cryptococcus neoformans* species complex. *Fungal Genet. Biol.* **78,** 16–48 (2015).
46. Cogliati, M. Global Molecular Epidemiology of Cryptococcus neoformans and Cryptococcus gattii: An Atlas of the Molecular Types. *Scientifica (Cairo).* **2013,** 675213 (2013).
47. Khawcharoenporn, T., Apisarnthanarak, A. & Mundy, L. M. Non- neoformans Cryptococcal Infections : a Systematic Review. *Infection* **35,** 51–58 (2007).
48. Bose, I., Reese, A., Ory, J., Doering, T. & Janbon, G. A Yeast under Cover : the Capsule of Cryptococcus neoformans. *Eukaryot. Cell* **2,** 655–663 (2003).
49. Araujo, G. D. S. *et al.* Capsules from Pathogenic and Non-Pathogenic Cryptococcus spp . Manifest Significant Differences in Structure and Ability to Protect against Phagocytic Cells. *PLoS One* **7,** (2012).
50. Ellis, D. H. & Pfeiffer, T. J. Natural Habitat of Cryptococcus neoformans var. gattii. *J. Clin. Microbiol.* **28,** 1642–1644 (1990).
51. Pfeiffer, T. & Ellis, D. Environmental Isolation of Cryptococcus neoformans gattii from California. *J. Immunol.* **163,** 929–930 (1991).
52. Springer, D. J. *et al.* Cryptococcus gattii VGIII Isolates Causing Infections in HIV / AIDS Patients in Southern California : Identification of the Local Environmental Source as Arboreal. *PLoS Pathog.* **10,** (2014).
53. Kidd, S. E. *et al.* Characterization of environmental sources of the human and animal pathogen Cryptococcus gattii in British Columbia, Canada, and the Pacific Northwest of the United States. *Appl. Environ. Microbiol.* **73,** 1433–43 (2007).
54. Lazéra, M. S. *et al.* Natural habitat of Cryptococcus neoformans var. neoformans in decaying wood forming hollows in living trees. *Med. Mycol.* **34,** 127–131 (1996).
55. Passoni, L. F. C., Wanke, B., Nishikawa, M. M. & Lazé Ra, M. S. Cryptococcus neoformans isolated from human dwellings in Rio de Janeiro, Brazil: an analysis of the domestic environment of AIDS patients with and without

cryptococcosis. *Med. Mycol.* **36,** 305–311 (1998).

56. Litvintseva, A. P. *et al.* Evidence that the human pathogenic fungus Cryptococcus neoformans var. grubii may have evolved in Africa. *PLoS One* **6,** e19688 (2011).

57. Kuroki, M. *et al.* Environmental isolation of Cryptococcus neoformans from an endemic region of HIV-associated cryptococcosis in Thailand. *Yeast* **21,** 809–812 (2004).

58. Currie, B. P., Freundlich, L. F. & Casadevalll, A. Restriction Fragment Length Polymorphism Analysis of Cryptococcus neoformans Isolates from Environmental (Pigeon Excreta) and Clinical Sources in New York City. *J. Clin. Microbiol.* **32,** 1188–1192 (1994).

59. Sriburee, P., Khayhan, S., Khamwan, C., Panjaisee, S. & Tharavichitkul, P. Serotype and PCR-fingerprints of clinical and environmental isolates of Cryptococcus neoformans in Chiang Mai, Thailand. *Mycopathologia* **158,** 25–31 (2004).

60. Fromtling, R. & Ruiz, A. Virulence and antifungal susceptibility of environmental and clinical isolates of Cryptococcus neoformans from Puerto Rico. *Mycopathologia* **160,** 163–166 (1989).

61. Da Silva, E. G. *et al.* Virulence profile of strains of Cryptococcus neoformans var. grubii evaluated by experimental infection in BALB/c mice and correlation with exoenzyme activity. *J. Med. Microbiol.* **55,** 139–142 (2006).

62. Litvintseva, A. P. & Mitchell, T. G. Most environmental isolates of Cryptococcus neoformans var. grubii (serotype A) are not lethal for mice. *Infect. Immun.* **77,** 3188–95 (2009).

63. Steele, K. T., Thakur, R., Nthobatsang, R., Steenhoff, A. P. & Bisson, G. P. In-hospital mortality of HIV-infected cryptococcal meningitis patients with *C. gattii* and *C. neoformans* infection in Gaborone, Botswana. *Med. Mycol.* **48,** 1112–1115 (2010).

64. Williamson, P. R. *et al.* Cryptococcal meningitis: epidemiology, immunology, diagnosis and therapy. *Nat. Rev. Neurol.* **13,** 13–24 (2016).

65. Chen, S. C. A. *et al.* Clinical manifestations of cryptococcus gattii infection:

Determinants of neurological sequelae and death. *Clin. Infect. Dis.* **55,** 789–798 (2012).

66. Beardsley, J. *et al.* Adjunctive Dexamethasone in HIV-Associated Cryptococcal Meningitis. *N. Engl. J. Med.* **374,** 542–554 (2016).
67. Day, J. N. *et al.* Combination antifungal therapy for cryptococcal meningitis. *N. Engl. J. Med.* **368,** 1291–302 (2013).
68. Wang, C.-Y. & Hsueh, P.-R. Nosocomial Transmission of Cryptococcosis. *N. Engl. J. Med.* 1271–1272 (2005).
69. Macdougall, L. & Fyfe, M. Emergence of Cryptococcus gattii in a Novel Environment Provides Clues to Its Incubation Period. *J. Clin. Microbiol.* **44,** 1851–1852 (2006).
70. Goldman, D. L. *et al.* Serologic Evidence for Cryptpcoccus neoformans infection in early childhood. *Pediatrics* **107,** (2001).
71. Saha, D. C. *et al.* Serologic Evidence for Reactivation of Cryptococcosis in Solid-Organ Transplant Recipients. *Clin. vaccine Immunol.* **14,** 1550–1554 (2007).
72. Garcia-Hermoso, D., Janbon, G. & Dromer, F. Epidemiological evidence for dormant Cryptococcus neoformans infection. *J. Clin. Microbiol.* **37,** 3204–9 (1999).
73. Lin, X. *et al.* Diploids in the Cryptococcus neoformans serotype A population homozygous for the anpha mating type originate via unisexual mating. *PLoS Pathog.* **5,** 10–15 (2009).
74. Heung, L. Innate Immune Responses to Cryptococcus. *J. Fungi* **3,** 35 (2017).
75. Esher, S. K., Zaragoza, O. & Alspaugh, J. A. Cryptococcal pathogenic mechanisms : a dangerous trip from the environment to the brain. **113,** 1–15 (2018).
76. Mukaremera, L. & Nielsen, K. Adaptive Immunity to Cryptococcus neoformans Infections. *J. Fungi* **3,** 1–20 (2017).
77. Rohatgi, S. & Pirofski, L. Host immunity to Cryptococcus neoformans. *Future Microbiol.* **10,** 565–581 (2015).
78. Beardsley, J. *et al.* Do Intracerebral Cytokine Responses Explain the Harmful Effects of Dexamethasone in Human Immunodeficiency Virus – associated

Cryptococcal Meningitis ? *Clin. Infect. Dis.* 1–8 (2018). doi:10.1093/cid/ciy725

79. Sloan, D. J. & Parris, V. Cryptococcal meningitis: Epidemiology and therapeutic options. *Clin. Epidemiol.* **6,** 169–182 (2014).

80. Jarvis, J. N. & Harrison, T. S. HIV-associated cryptococcal meningitis. *AIDS* **21,** 2119–2129 (2007).

81. Chariyalertsak, S., Sirisanthana, T., Saengwonloey, O. & Nelson, K. E. Clinical presentation and risk behaviors of patients with acquired immunodeficiency syndrome in Thailand, 1994--1998: regional variation and temporal trends. *Clin. Infect. Dis.* **32,** 955–62 (2001).

82. Jarvis, J. N. *et al.* Determinants of mortality in a combined cohort of 501 patients with HIV-associated cryptococcal meningitis: Implications for improving outcomes. *Clin. Infect. Dis.* **58,** 736–745 (2014).

83. Neofytos, D. *et al.* Epidemiology and outcome of invasive fungal infections in solid organ transplant recipients. *Transpl. Infect. Dis.* **12,** 220–229 (2010).

84. Pappas, P. G. Cryptococcal infections in non-HIV-infected patients. *Trans. Am. Clin. Climatol. Assoc.* **124,** 61–79 (2013).

85. Brizendine, K. D., Baddley, J. W. & Pappas, P. G. Predictors of Mortality and Differences in Clinical Features among Patients with Cryptococcosis According to Immune Status. *PLoS One* **8,** (2013).

86. Day, J. N. *et al.* Most cases of cryptococcal meningitis in HIV-uninfected patients in Vietnam are due to a distinct amplified fragment length polymorphism-defined cluster of Cryptococcus neoformans var. grubii VN1. *J. Clin. Microbiol.* **49,** 658–64 (2011).

87. Chen, J. *et al. Cryptococcus neoformans* strains and infection in apparently immunocompetent patients, China. *Emerg. Infect. Dis.* **14,** 755–762 (2008).

88. Day, J. N. *et al.* Comparative genomics of Cryptococcus neoformans var. grubii associated with meningitis in HIV infected and uninfected patients in Vietnam. *PLoS Negl. Trop. Dis.* **11,** 1–22 (2017).

89. Thompson, G. R. *et al.* Antifungal susceptibilities among different serotypes of Cryptococcus gattii and Cryptococcus neoformans. *Antimicrob. Agents Chemother.* **53,** 309–11 (2009).

90. Kronstad, J. W. *et al.* Expanding fungal pathogenesis: Cryptococcus breaks out of the opportunistic box. *Nat. Rev. Microbiol.* **9,** 193–203 (2011).
91. Steenbergen, J. N. & Casadevall, A. The origin and maintenance of virulence for the human pathogenic fungus Cryptococcus neoformans. *Microbes Infect.* **5,** 667–75 (2003).
92. Perfect, J. R. Cryptococcus neoformans: The yeast that likes it hot. *FEMS Yeast Res.* **6,** 463–468 (2006).
93. Petter, R., Kang, B. S., Boekhout, T., Davis, B. J. & Kwon-Chung, K. J. A survey of heterobasidiomycetous yeasts for the presence of the genes homologous to virulence factors of Filobasidiella neoformans, CNLAC1 and CAP59. *Microbiology* **147,** 2029–2036 (2001).
94. Odom, a *et al.* Calcineurin is required for virulence of Cryptococcus neoformans. *EMBO J.* **16,** 2576–2589 (1997).
95. Fan, W., Kraus, P. R., Boily, M. & Heitman, J. Cryptococcus neoformans Gene Expression during Murine Macrophage Infection. *Eukaryot. Cell* **4,** 1420–1433 (2005).
96. Hu, G. *et al.* PI3K signaling of autophagy is required for starvation tolerance and virulence of C. neoformans. *J. Clin. Invest.* **118,** (2008).
97. Chang, Y. C. & Kwon-Chung, K. J. Complementation of a capsule-deficient mutation of Cryptococcus neoformans restores its virulence. *Mol. Cell. Biol.* **14,** 4912–9 (1994).
98. Mylonakis, E. *et al. Galleria mellonella* as a model system to study *Cryptococcus neoformans* pathogenesis. *Infect. Immun.* **73,** 3842–2850 (2005).
99. Zaragoza, O., Frases, S. & Casadevalll, A. Capsule enlargement in Cryptococcus neoformans confers resistance to oxidative stress suggesting a mechanism for intracellular survival. *Cell Microbiol.* **10,** 2043–2057
100. Zaragoza, O. *et al. The Capsule of the Fungal Pathogen Cryptococcus neoformans*. *Advances in Applied Microbiology* **68,** (Elsevier Inc., 2009).
101. Jain, N. *et al.* Phenotypic Switching in a Cryptococcus neoformans Variety gattii Strain Is Associated with Changes in Virulence and Promotes Dissemination to the Central Nervous System. *Infect. Immun.* **74,** 896–903

(2006).

102. Tucker, S. C. & Casadevall, A. Replication of Cryptococcus neoformans in macrophages is accompanied by phagosomal permeabilization and accumulation of vesicles containing polysaccharide in the cytoplasm. *Proc. Natl. Acad. Sci.* **99,** 3165–3170 (2002).
103. Chang, Y. C. & Kwon-Chung, K. J. Isolation, characterization, and localization of a capsule-associated gene, CAP10, of Cryptococcus neoformans. *J. Bacteriol.* **181,** 5636–5643 (1999).
104. Eisenman, H. C. *et al.* Microstructure of cell wall-associated melanin in the human pathogenic fungus Cryptococcus neoformans. *Biochemistry* **44,** 3683–3693 (2005).
105. Liu, G. Y. & Nizet, V. Color me bad: microbial pigments as virulence factors. *Trends Microbiol.* **17,** 406–413 (2009).
106. Duin, D. Van, Casadevall, A. & Nosanchuk, J. D. Melanization of Cryptococcus neoformans and Histoplasma capsulatum reduces their susceptibilities to Amphotericin B and Caspofungin. *Antimicrob. Agents Chemother.* **46,** 3394–3400 (2002).
107. Salas, S., Perfect, J. & Williamson, P. Effect of Laccase gene, CNLAC1, on virulence of Cryptococcus neoformans. *J. Exp. Med.* **184,** 377–386 (1996).
108. De Jesús-Berríos, M. *et al.* Enzymes that Counteract Nitrosative Stress Promote Fungal Virulence. *Curr. Biol.* **13,** 1963–1968 (2003).
109. Seider, K., Heyken, A., Lüttich, A., Miramón, P. & Hube, B. Interaction of pathogenic yeasts with phagocytes: Survival, persistence and escape. *Curr. Opin. Microbiol.* **13,** 392–400 (2010).
110. O'Meara, T. R. *et al.* Interaction of Cryptococcus neoformans Rim101 and protein kinase a regulates capsule. *PLoS Pathog.* **6,** (2010).
111. Meara, T. R. O., Holmer, S. M., Selvig, K., Dietrich, F. & Alspaugh, J. A. Cryptococcus neoformans Rim101 Is Associated with Cell Wall. *MBio* **4,** 1–13 (2013).
112. Qiu, J., Olszewski, M. A. & Williamson, P. R. Cryptococcus neoformans growth and protection from innate immunity are dependent on expression of a DEAD

box protein VAD1. *Infect. Immun.* **81,** (2013).

113. Luberto, C. *et al.* Identification of App1 as a regulator of phagocytosis and virulence of Cryptococcus neoformans. *J. Clin. Invest.* **112,** (2003).
114. Stano, P. *et al.* App1: An Antiphagocytic Protein That Binds to Complement Receptors 3 and 2. *J. Immunol.* **182,** (2009).
115. Liu, O. W. *et al.* Systematic Genetic Analysis of Virulence in the Human Fungal Pathogen Cryptococcus neoformans. *Cell* **135,** 174–188 (2008).
116. Chun, C. D., Brown, J. C. S. & Madhani, H. D. A major role for capsule-independent phagocytosis-inhibitory mechanisms in mammalian infection by Cryptococcus neoformans. *Cell Host Microbe* **9,** 243–251 (2011).
117. Rodrigues, M. L. *et al.* Vesicular polysaccharide export in Cryptococcus neoformans is a eukaryotic solution to the problem of fungal trans-cell wall transport. *Eukaryot. Cell* **6,** 48–59 (2007).
118. Monari, C., Bistoni, F. & Vecchiarelli, A. Glucuronoxylomannan exhibits potent immunosuppressive properties. *FEMS Yeast Res.* **6,** 537–542 (2006).
119. Rodrigues, M. L. *et al.* Extracellular vesicles produced by Cryptococcus neoformans contain protein components associated with virulence. *Eukaryot. Cell* **7,** 58–67 (2008).
120. D'Souza, C. A. & Heitman, J. It infects me, it infects me not: Phenotypic switching in the fungal pathogen Cryptococcus neoformans. *J. Clin. Invest.* **108,** 1577–1578 (2001).
121. Casadevall, A., Goldman, D., Fries, B. & Montella, L. Phenotypic switching in the human pathogenic fungus Cryptococcus neoformans is associated with changes in virulence and pulmonary inflammatory response in rodents. *Proc. Natl. Acad. Sci.* **95,** 14967–14972 (1998).
122. Fries, B. C., Taborda, C. P., Serfass, E. & Casadevall, a. Phenotypic switching of Cryptococcus neoformans occurs in vivo and influences the outcome of infection. *J. Clin. Invest.* **108,** 1639–1648 (2001).
123. Jain, N., Guerrero, A. & Fries, B. C. Phenotypic switching and its implications for the pathogenesis of Cryptococcus neoformans. *FEMS Yeast Res.* **6,** 480–488 (2006).

124. Neilson, J. B., Ivey, M. H. & Bulmer, G. S. Cryptococcus neoformans: pseudohyphal forms surviving culture with Acanthamoeba polyphaga. *Infect. Immun.* **20,** 262–266 (1978).

125. Okagaki, L. H. *et al.* Cryptococcal cell morphology affects host cell interactions and pathogenicity. *PLoS Pathog.* **6,** (2010).

126. Okagaki, L. H. & Nielsen, K. Titan Cells Confer Protection from Phagocytosis in Cryptococcus neoformans Infections. **11,** 820–826 (2012).

127. Zaragoza, O. Fungal Cell Gigantism during Mammalian Infection. *PLoS Pathog.* **6,** (2010).

128. Zaragoza, O. & Nielsen, K. Titan cells in Cryptococcus neoformans : cells with a giant impact. *Curr. Opin. Microbiol.* **16,** 409–413 (2013).

129. Chen, L. C., Blank, E. S. & Casadevall, A. Extracellular proteinase activity of Cryptococcus neoformans. *Clin. Diagn. Lab. Immunol.* **3,** 570–574 (1996).

130. Chrisman, C. J., Albuquerque, P., Guimaraes, A. J., Nieves, E. & Casadevall, A. Phospholipids trigger Cryptococcus neoformans capsular enlargement during interactions with amoebae and macrophages. *PLoS Pathog.* **7,** (2011).

131. Cox, G. M. *et al.* Extracellular phospholipase activity is a virulence factor for Cryptococcus neoformans. *Mol. Microbiol.* **39,** 166–175 (2001).

132. Kronstad, J. W. *et al.* Expanding fungal pathogenesis : Cryptococcus breaks out of the opportunistic box. *Nat. Rev. Microbiol.* **9,** (2011).

133. Cox, G. M., Mukherjee, J., Cole, G. T., Casadevall, a & Perfect, J. R. Urease as a virulence factor in experimental cryptococcosis. *Infect. Immun.* **68,** 443–8 (2000).

134. Varma, A. *et al.* Identification of a novel gene , URE2 , that functionally complements a urease-negative clinical strain of Cryptococcus neoformans. *Microbiology* **152,** 3723–3731 (2006).

135. Zimmer, B. L. & Roberts, G. D. Rapid selective urease test for presumptive identification of Cryptococcus neoformans. *J. Clin. Microbiol.* **10,** 380–381 (1979).

136. Mains, P. E., Sulston, I. A. & Wood, W. B. G protein-coupled receptor Gpr4 senses amino acids and activates the cAMP-PKA pathway in Cryptococcus

neoformans. *Mol. Biol. Cell* **17,** 667–679 (2006).

137. Bahn, Y. S., Hicks, J. K., Giles, S. S., Cox, G. M. & Heitman, J. Adenylyl cyclase-associated protein Aca1 regulates virulence and differentiation of Cryptococcus neoformans via the cyclic AMP-protein kinase A cascade. *Eukaryot. Cell* **3,** 1476–1491 (2004).
138. Kozubowski, L., Lee, S. C. & Heitman, J. Signalling pathways in the pathogenesis of Cryptococcus. *Cell. Microbiol.* **11,** 370–380 (2009).
139. Kronstad, J. W., Hu, G. & Choi, J. The cAMP / Protein Kinase A Pathway and Virulence in Cryptococcus neoformans. *Mycobiology* **39,** 143–150 (2011).
140. Hu, G. *et al.* Transcriptional Regulation by Protein Kinase A in Cryptococcus neoformans. *PLoS Pathog.* **3,** (2007).
141. Maeng, S. *et al.* Comparative transcriptome analysis reveals novel roles of the Ras and cyclic AMP signaling pathways in environmental stress response and antifungal drug sensitivity in Cryptococcus neoformans. *Eukaryot Cell* **9,** 360–378 (2010).
142. Román, E., Arana, D. M., Nombela, C., Alonso-Monge, R. & Pla, J. MAP kinase pathways as regulators of fungal virulence. *Trends Microbiol.* **15,** 181–190 (2007).
143. O'Meara, T. R. & Andrew Alspaugh, J. The Cryptococcus neoformans capsule: A sword and a shield. *Clin. Microbiol. Rev.* **25,** 387–408 (2012).
144. Bahn, Y. S., Kojima, K., Cox, G. M. & Heitman, J. A Unique Fungal Two component system regulate stress responses, drug sensitivity, sexual development and virulence of C. neoformans. *Mol. Biol. Cell* **17,** 3122–3135 (2006).
145. Bahn, Y.-S., Cox, G. M. & Heitman, J. Specialization of the HOG pathway and its impact on differentiation and virulence of Crypto. *Mol. Biol. Cell* **16,** 1–13 (2005).
146. Gerik, K. J. *et al.* PKC1 Is Essential for Protection against both Oxidative and Nitrosative Stresses , Cell Integrity , and Normal Manifestation of Virulence Factors in the Pathogenic Fungus Cryptococcus neoformans. *Eukaryot. Cell* **7,** 1685–1698 (2008).

147. Gerik, K. J. *et al.* Cell wall integrity is dependent on the PKC1 signal transduction pathway in Cryptococcus neoformans. *Mol. Microbiol.* **58,** 393–408 (2005).
148. Kraus, P. R., Fox, D. S., Cox, G. M. & Heitman, J. The Cryptococcus neoformans MAP kinase Mpk1 regulates cell integrity in response to antifungal drugs and loss of calcineurin function. *Mol. Microbiol.* **48,** 1377–1387 (2003).
149. Wang, Y., Liu, T. B., Patel, S., Jiang, L. & Xue, C. The casein kinase i protein cck1 regulates multiple signaling pathways and is essential for cell integrity and fungal virulence in cryptococcus neoformans. *Eukaryot. Cell* **10,** 1455–1464 (2011).
150. Bahn, Y. S. & Jung, K. W. Stress signaling pathways for the pathogenicity of Cryptococcus. *Eukaryot. Cell* **12,** 1564–1577 (2013).
151. Odom, A. *et al.* Calcineurin is required for virulence of Cryptococcus neoformans. *EMBO J.* **16,** 2576–2589 (1997).
152. Kraus, P. R., Nichols, C. B. & Heitman, J. Calcium- and calcineurin-independent roles for calmodulin in Cryptococcus neoformans morphogenesis and high-temperature growth. *Eukaryot. Cell* **4,** 1079–1087 (2005).
153. Fox, D. S. *et al.* Calcineurin regulatory subunit is essential for virulence and mediates interactions with FKBP12 - FK506 in Cryptococcus neoformans. *Mol. Microbiol.* **39,** (2001).
154. Cruz, M. C. & Fox, D. S. Calcineurin is required for hyphal elongation during mating and haploid fruiting in Cryptococcus neoformans. *EMBO J.* **20,** (2001).
155. Souza, C. A. D. *et al.* Cyclic AMP-Dependent Protein Kinase Controls Virulence of the Fungal Pathogen Cryptococcus neoformans Cyclic AMP-Dependent Protein Kinase Controls Virulence of the Fungal Pathogen Cryptococcus neoformans. *Mol. Cell. Biol.* **21,** 3179–3191 (2001).
156. Alspaugh, J. A., Cavallo, L. M., Perfect, J. R. & Heitman, J. RAS1 regulates filamentation, mating and growth at high temperature of Cryptococcus neoformans. *Mol. Microbiol.* **36,** 352–365 (2000).
157. Waugh, M. S., Vallim, M. A., Heitman, J. & Alspaugh, J. A. Ras1 controls pheromone expression and response during mating in Cryptococcus

neoformans. *Fungal Genet. Biol.* **38,** 110–121 (2003).

158. Waugh, M. S. *et al.* Ras1 and Ras2 contribute shared and unique roles in physiology and virulence of Cryptococcus neoformans. *Microbiology* **148,** 191–201 (2002).

159. Nichols, C. B., Perfect, Z. H. & Alspaugh, J. A. A Ras1-Cdc24 signal transduction pathway mediates thermotolerance in the fungal pathogen Cryptococcus neoformans. *Mol. Microbiol.* **63,** 1118–1130 (2007).

160. Sabiiti, W., May, R. C. & Pursall, E. R. Experimental models of cryptococcosis. *Int. J. Microbiol.* **2012,** (2012).

161. Mylonakis, E., Casadevall, A. & Ausubel, F. M. Exploiting amoeboid and non-vertebrate animal model systems to study the virulence of human pathogenic fungi. *PLoS Pathog.* **3,** 0859–0865 (2007).

162. Cook, S. M. & McArthur, J. D. Developing Galleria mellonella as a model host for human pathogens. *Virulence* **4,** 350–3 (2013).

163. Jander, G., Rahme, L. G. & Frederick, M. Positive Correlation between Virulence of Pseudomonas aeruginosa Mutants in Mice and Insects. *J. Bacteriol.* **182,** 3843–3845 (2000).

164. Brennan, M., Thomas, D. Y., Whiteway, M. & Kavanagh, K. Correlation between virulence of *Candida albicans* mutants in mice and *Galleria mellonella* larvae. *FEMS Immunol. Med. Microbiol.* **34,** 153–157 (2002).

165. Kavanagh, K. & Fallon, J. P. Galleria mellonella larvae as models for studying fungal virulence. *Fungal Biol. Rev.* **24,** 79–83 (2010).

166. Firacative, C., Duan, S. & Meyer, W. Galleria mellonella model identifies highly virulent strains among all major molecular types of Cryptococcus gattii. *PLoS One* **9,** (2014).

167. García-Rodas, R., Casadevall, A., Rodríguez-Tudela, J. L., Cuenca-Estrella, M. & Zaragoza, O. Cryptococcus neoformans capsular enlargement and cellular gigantism during Galleria mellonella infection. *PLoS One* **6,** (2011).

168. Banville, N., Browne, N. & Kavanagh, K. Effect of nutrient deprivation on the susceptibility of Galleria mellonella larvae to infection. *Virulence* **3,** 497–503 (2012).

169. Chen, Y.-H. *et al.* Multilocus Sequence Typing Reveals both Shared and Unique Genotypes of Cryptococcus neoformans in Jiangxi Province, China. *Sci. Rep.* **8,** 1495 (2018).
170. Mak, S., Klinkenberg, B., Bartlett, K. & Fyfe, M. Ecological niche modeling of cryptococcus gattii in British Columbia, Canada. *Environ. Health Perspect.* **118,** 653–658 (2010).
171. Litvintseva, A. P., Kestenbaum, L., Vilgalys, R. & Mitchell, T. G. Comparative Analysis of Environmental and Clinical Populations of Cryptococcus neoformans. *J. Clin. Microbiol.* **43,** (2005).
172. Granados, D. P. & Castañeda, E. Influence of climatic conditions on the isolation of members of the Cryptococcus neoformans species complex from trees in Colombia from 1992-2004. *FEMS Yeast Res.* **6,** 636–644 (2006).
173. Vanhove, M. *et al.* Genomic epidemiology of Cryptococcus yeasts identifies adaptation to environmental niches underpinning infection across an African HIV/AIDS cohort. *Mol. Ecol.* **26,** 1991–2005 (2017).
174. Coleman, M. *et al.* Using the SaTScan method to detect local malaria clusters for guiding malaria control programmes. *Malar. J.* **8,** (2009).
175. Mostashari M & Miller, J, F. K. Dead bird clustering as an early warning system for West Nile virus activity. *Emerg. Infect. Dis.* **9,** 641–646 (2003).
176. Wu, W., Guo, J., Guan, P., Sun, Y. & Zhou, B. Clusters of spatial, temporal, and space-time distribution of hemorrhagic fever with renal syndrome in Liaoning Province, Northeastern China. *BMC Infect. Dis.* **11,** (2011).
177. Wheeler, D. C. A comparison of spatial clustering and cluster detection techniques for childhood leukemia incidence in Ohio, 1996-2003. *Int. J. Health Geogr.* **6,** 1–16 (2007).
178. Deng, T. *et al.* Spatial-Temporal Clusters and Risk Factors of Hand, Foot, and Mouth Disease at the District Level in Guangdong Province, China. *PLoS One* **8,** (2013).
179. Pham Thanh, D. *et al.* The Molecular and Spatial Epidemiology of Typhoid Fever in Rural Cambodia. *PLoS Negl. Trop. Dis.* **10,** 1–16 (2016).
180. Chau, T. T. *et al.* A prospective descriptive study of cryptococcal meningitis in

HIV uninfected patients in Vietnam - high prevalence of Cryptococcus neoformans var grubii in the absence of underlying disease. *BMC Infect. Dis.* **10,** 199 (2010).

181. Kulldorff, M. A spatial scan statistics. *Commun. Stat. - Theory Methods* **26,** (1997).
182. Ho Dang Trung, N. *et al.* Aetiologies of central nervous system infection in Viet Nam: A prospective provincial hospital-based descriptive surveillance study. *PLoS One* **7,** (2012).
183. Smith, R., Khan, M. & Dao, V. *HIV/AIDS in Viet Nam: The current situation, The national response, the emerging challenges*. (2006).
184. Wiesner, D. L. *et al.* Cryptococcal Genotype Influences Immunologic Response and Human. **3,** 1–10 (2012).
185. Coleman, M. *et al.* Using the SaTScan method to detect local malaria clusters for guiding malaria control programmes. *Malar. J.* **8,** 68 (2009).
186. Greene, S. K., Peterson, E. R., Kapell, D., Fine, A. D. & Kulldorff, M. Daily Reportable Disease Spatiotemporal Cluster Detection, New York City, New York, USA, 2014-2015. *Emerg. Infect. Dis.* **22,** 1808–1812 (2016).
187. Baker, S. *et al.* Combined high-resolution genotyping and geospatial analysis reveals modes of endemic urban typhoid fever transmission. *Open Biol.* **1,** 110008–110008 (2011).
188. Aamodt, G., Samuelsen, S. O. & Skrondal, A. A simulation study of three methods for detecting disease clusters. *Int. J. Health Geogr.* **5,** 15 (2006).
189. Tango, T. & Takahashi, K. A flexibly shaped spatial scan statistic for detecting clusters. *Int. J. Health Geogr.* **4,** 1–15 (2005).
190. Yamamoto, Y. *et al.* Random amplified polymorphic DNA analysis of clinically and environmentally isolated Cryptococcus neoformans in Nagasaki. *J. Clin. Microbiol.* **33,** 3328–3332 (1995).
191. Sorrell, T. C. *et al.* Concordance of clinical and environmental isolates of Cryptococcus neoformans var. gattii by random amplification of polymorphic DNA analysis and PCR fingerprinting. *J. Clin. Microbiol.* **34,** 1253–1260 (1996).
192. Randhawa, H. S. *et al.* Seasonal variations in the prevalence of Cryptococcus

neoformans var. grubii and Cryptococcus gattii in decayed wood inside trunk hollows of diverse tree species in north-western India: a retrospective study. *Med. Mycol.* **49,** 320–323 (2011).

193. Illnait-Zaragozí, M. T. *et al.* Environmental isolation and characterisation of Cryptococcus species from living trees in Havana city, Cuba. *Mycoses* **55,** (2012).
194. Granados, D. P. & Castañeda, E. Isolation and characterization of Cryptococcus neoformans varieties recovered from natural sources in Bogotá, Colombia, and study of ecological conditions in the area. *Microb. Ecol.* **49,** 282–90 (2005).
195. Le, T. *et al.* Epidemiology, seasonality, and predictors of outcome of aids-associated penicillium marneffei infection in Ho Chi Minh City, Viet Nam. *Clin. Infect. Dis.* **52,** 945–952 (2011).
196. Pham, C. D., Ahn, S., Turner, L. a, Wohrle, R. & Lockhart, S. R. Development and validation of benomyl birdseed agar for the isolation of Cryptococcus neoformans and Cryptococcus gattii from environmental samples. *Med. Mycol.* **52,** 417–21 (2014).
197. Shileds, A. & Ajello, L. Medium for selective isolation of Cryptococcus neoformans. *Science (80-.).* **151,** 208–209 (1966).
198. Viviani, M. A., Wen, H. & Tortorano, A. A simplified method to extract high-quality DNA from Cryptococcus neoformans. *J. Mycol. Med* **6,** (1996).
199. Begerow, D., Nilsson, H., Unterseher, M. & Maier, W. Current state and perspectives of fungal DNA barcoding and rapid identification procedures. *Appl. Microbiol. Biotechnol.* **87,** 99–108 (2010).
200. White, T., Bruns, T. & Taylor, J. *Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics*. (Academic press, 1990).
201. Lee, A. *et al.* Survival defects of Cryptococcus neoformans mutants exposed to human cerebrospinal fluid result in attenuated virulence in an experimental model of meningitis. *Infect. Immun.* **78,** 4213–4225 (2010).
202. Thompson, J., Higgins, D. & Gibson, T. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting,

position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22,** (1994).

203. Saitou, N. & Nei, M. The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol Biol Evol* **4,** (1987).
204. Granados, D. P. & Castañeda, E. Isolation and characterization of Cryptococcus neoformans varieties recovered from natural sources in Bogotá, Colombia, and study of ecological conditions in the area. *Microb. Ecol.* **49,** 282–90 (2005).
205. Caicedo, L. D., Alvarez, M. I., Delgado, M. & Cárdenas, A. Cryptococcus neoformans in bird excreta in the city zoo of Cali, Colombia. *Mycopathologia* **147,** 121–124 (1999).
206. Pereira, C. B. *et al.* Evaluation of laccases and melanization in clinical and environmental Cryptococcus neoformans samples by non-denaturing PAGE. *J. Med. Microbiol.* **58,** 563–566 (2009).
207. Bo, E. C., Altwegg, M. & Bosshard, P. P. Internal Transcribed Spacer Sequencing versus Biochemical Profiling for Identification of Medically Important Yeasts. **44,** 77–84 (2006).
208. Duran-Valle, T., Sanz-Rodriguez, N. & Gomez-Garces, J. Identification of clinical yeasts by Vitek MS system compared with API ID 32 C. *Med. Mycol.* **52,** 342–349 (2014).
209. Loeffler, J. *et al.* Contamination Ocurring in Fungal PCR Assays. *J. Clin. Microbiol.* **37,** 1200–1202 (1999).
210. Czurda, S., Smelik, S., Preuner-Stix, S., Nogueira, F. & Lion, T. Occurrence of fungal DNA contamination in PCR reagents: Approaches to control and decontamination. *J. Clin. Microbiol.* **54,** 148–152 (2016).
211. Smith, N., Sehring, M., Chambers, J. & Patel, P. Perspectives on non-*neoformans* cryptococcal opportunistic infections. *J. Community Hosp. Intern. Med. Perspect.* **7,** 214–217 (2017).
212. Nakase, T. & Tanticharoen, M. *Yeast Biodiversity in Tropical Forests of Asia*. *Biodiversity and Ecophysiology of Yeasts* (2006).
213. de Garcia, V., Zalar, P., Brizzio, S., Gunde-Cimerman, N. & van Broock, M.

Cryptococcus species (Tremellales) from glacial biomes in the southern (Patagonia) and northern (Svalbard) hemispheres. *FEMS Microbiol. Ecol.* **82,** 523–539 (2012).

214. Saluja, P. & Prasad, G. S. Crytococcus rajasthanensis sp. nov., an anamorphic yeast species related to Cryptococcus laurentii, isolated from Rajasthan, India. *Int. J. Syst. Evol. Microbiol.* **57,** 414–418 (2007).
215. Ferreira-Paim, K. *et al.* Phylogenetic Analysis of Phenotypically Characterized Cryptococcus laurentii Isolates Reveals High Frequency of Cryptic Species. *PLoS One* **9,** e108633 (2014).
216. Randhawa, H. S. *et al.* The expanding host tree species spectrum of Cryptococcus gattii and Cryptococcus neoformans and their isolations from surrounding soil in India. *Med. Mycol.* **46,** 823–833 (2008).
217. McTaggart, A. R., Doungsa-ard, C., Geering, A. D. W., Aime, M. C. & Shivas, R. G. A co-evolutionary relationship exists between Endoraecium (Pucciniales) and its Acacia hosts in Australia. *Persoonia - Mol. Phylogeny Evol. Fungi* **35,** 50–62 (2015).
218. Chen, Y. *et al.* Comparative analyses of clinical and environmental populations of Cryptococcus neoformans in Botswana. *Mol. Ecol.* **24,** 3559–3571 (2015).
219. Vanhove, M. & Mathieu. Genomic epidemiology and ecology of cryptococcus neoformans var. grubii in southern Africa. (Imperial College London, 2016). at <https://spiral.imperial.ac.uk/handle/10044/1/42540>
220. Conde-Pereira, C. *et al.* Fatal case of polymicrobial meningitis caused by Cryptococcus liquefaciens and Mycobacterium tuberculosis complex in a human immunodeficiency virus-infected patient. *J. Clin. Microbiol.* **53,** 2753–2755 (2015).
221. De Araújo, G. R. S. *et al.* The environmental yeast Cryptococcus liquefaciens produces capsular and secreted polysaccharides with similar pathogenic properties to those of C. neoformans. *Sci. Rep.* **7,** 1–12 (2017).
222. Garcia-solache, M. A. & Casadevall, A. Diseases for Mammals Global Warming Will Bring New Fungal Diseases for Mammals. **1,** 1–3 (2010).
223. Fang, W., Fa, Z. & Liao, W. Epidemiology of Cryptococcus and cryptococcosis in

China. *Fungal Genet. Biol.* **78,** 7–15 (2015).

224. Mihara, T. *et al.* Multilocus sequence typing of Cryptococcus neoformans in non-HIV associated cryptococcosis in Nagasaki, Japan. *Med. Mycol.* **51,** 252–260 (2013).
225. Binder, U., Maurer, E. & Lass-Flörl, C. Galleria mellonella: An invertebrate model to study pathogenicity in correctly defined fungal species. *Fungal Biol.* **120,** 288–295 (2016).
226. Liu, O. W. *et al.* Systematic Genetic Analysis of Virulence in the Human Fungal Pathogen Cryptococcus neoformans. *Cell* **135,** 174–188 (2008).
227. Lee, K. T. *et al.* Systematic functional analysis of kinases in the fungal pathogen Cryptococcus neoformans. *Nat. Commun.* **7,** 1–16 (2016).
228. Crabtree, J. N. *et al.* Titan Cell Production Enhances the Virulence of Cryptococcus neoformans. *Infect. Immun.* **80,** 3776–3785 (2012).
229. Kavanagh, K. & Reeves, E. P. Exploiting the potential of insects for in vivo pathogenicity testing of microbial pathogens. *FEMS Microbiol. Rev.* **28,** 101–112 (2004).
230. Bielska, E. *et al.* Pathogen-derived extracellular vesicles mediate virulence in the fatal human pathogen Cryptococcus gattii. *Nat. Commun.* **9,** (2018).
231. Lortholary, O. *et al.* Cytokine Profiles of AIDS Patients Are Similar to Those of Mice with Disseminated Cryptococcus neoformans Infection Cytokine Profiles of AIDS Patients Are Similar to Those of Mice with Disseminated Cryptococcus neoformans Infection. (1999).
232. Bergin, D. *et al.* Superoxide Production in Galleria mellonella Hemocytes : Identification of Proteins Homologous to the NADPH Oxidase Complex of Human Neutrophils Superoxide Production in Galleria mellonella Hemocytes : Identification of Proteins Homologous to the NADPH Ox. *Infect Immun* **73,** 4161–4170 (2005).
233. Voelz, K. & May, R. C. Cryptococcal interactions with the host immune system. *Eukaryot. Cell* **9,** 835–846 (2010).
234. Ngamskulrungroj, P., Serena, C., Gilgado, F., Malik, R. & Meyer, W. Global VGIIa isolates are of comparable virulence to the major fatal Cryptococcus

gattii Vancouver Island outbreak genotype. *Clin. Microbiol. Infect.* **17,** 251–258 (2011).

235. Rhodes, J. *et al.* A Population Genomics Approach to Assessing the Genetic Basis of Within-Host Microevolution Underlying Recurrent Cryptococcal Meningitis Infection. *G3* **7,** 1165–1176 (2017).
236. Cheema, M. S. & Christians, J. K. Virulence in an insect model differs between mating types in Aspergillus fumigatus. *Med. Mycol.* **49,** 202–207 (2011).
237. Gugnani, H. *et al.* Role of Cannomys badius as a Natural Animal Host of Penicillium marneffei in India. **42,** 5070–5075 (2004).
238. Junior, A., Santos, B., Carvalho, E. & Stoianoff, M. Biological activity of Cryptococcus neoformans and Cryptococcus gattii from clinical and environmental isolates. *J Bras Patol Med Lab* **49,** 160–168 (2013).
239. Capilla, J., Maffei, C. M. L., Clemons, K. V., Sobel, R. A. & Stevens, D. A. Experimental systemic infection with Cryptococcus neoformans var. grubii and Cryptococcus gattii in normal and immunodeficient mice. *Med. Mycol.* **44,** 601–610 (2006).
240. Arras, S. D. M. *et al.* Convergent microevolution of Cryptococcus neoformans hypervirulence in the laboratory and the clinic. *Sci. Rep.* **7,** 1–14 (2017).
241. McClelland, E. E., Perrine, W. T., Potts, W. K. & Casadevall, A. Relationship of virulence factor expression to evolved virulence in mouse-passaged Cryptococcus neoformans lines. *Infect. Immun.* **73,** 7047–7050 (2005).
242. Albuquerque, P. & Casadevall, A. Quorum sensing in fungi a review. *Med. Mycol.* **50,** 337–345 (2012).
243. Rutherford, S. T. & Bassler, B. L. Bacterial Quorum Sensing : Its Role in Virulence and Possibilities for Its Control. *Cold Spring Harb. Perspect. Med.* **2,** 1–26 (2012).
244. Hornby, J. M. *et al.* Quorum Sensing in the Dimorphic Fungus Candida albicans Is Mediated by Farnesol. **67,** 2982–2992 (2001).
245. Lee, H., Chang, Y. C., Nardone, G. & Kwon-Chung, K. J. TUP1 disruption in Cryptococcus neoformans uncovers a peptide-mediated density-dependent growth phenomenon that mimics quorum sensing. *Mol. Microbiol.* **64,** 591–

601 (2007).

246. Homer, C. M. *et al.* Intracellular Action of a Secreted Peptide Required for Fungal Virulence. *Cell Host Microbe* **19,** 849–864 (2016).

247. Thanh, L. T. Characterization of C . neoformans var . grubii Isolates From HIV-Infected and HIV-uninfected patients from Vietnam. (Open University, 2018).

248. Bodi, Z., Farkas, Z., Papp, B., Moura, G. & Pal, C. Phenotypic heterogeneity promotes adaptive evolution. *PLoS Biol.* **15,** (2017).

249. Son, M., Yu, J. & Kim, K. H. Five Questions about Mycoviruses. *PLoS Pathog.* **11,** 5–11 (2015).

250. Halfmann, R. *et al.* Prions are a common mechanism for phenotypic inheritance in wild yeasts. *Nature* **482,** 363–368 (2012).

251. Liebman, S. W. & Chernoff, Y. O. Prions in yeast. *Genetics* **191,** 1041–1072 (2012).

252. Kent, C. R., Ortiz-bermu, P., Giles, S. S. & Hull, C. M. Formulation of a Defined V8 Medium for Induction of Sexual Development of Cryptococcus neoformans . *Appl. Environ. Microbiol.* **74,** 6248–6253 (2008).

253. Neilson, J. B., Fromtling, R. A. & Bulmer, G. S. Pseudohyphal forms of Cryptococcus neoformans: decreased survival in vivo. *Mycopathologia* **59,** 57–59 (1981).

254. Feldmesser, M., Kress, Y. & Casadevall, A. Dynamic changes in the morphology of Cryptococcus neoformans during murine pulmonary infection. *Microbiology* **147,** 2355–2365 (2001).

255. Steen, B. R. *et al.* Cryptococcus neoformans gene expression during experimental cryptococcal meningitis. *Eukaryot Cell* **2,** 1336–1349 (2003).

256. Lorenz, M. C., Bender, J. a & Fink, G. R. Transcriptional Response of Candida albicans upon Internalization by Macrophages Transcriptional Response of Candida albicans upon Internalization by Macrophages. *Eukaryot. Cell* **3,** 1076–1087 (2004).

257. Kukurba, K. R. & Montgomery, S. B. RNA sequencing and analysis. *Cold Spring Harb. Protoc.* **2015,** 951–969 (2015).

258. Wang, Z., Gerstein, M. & Snyder, M. RNA-Seq : a revolutionary tool for

transcriptomics. *Nat. Rev. Genet.* **10,** (2009).

259. Janbon, G. *et al.* Analysis of the Genome and Transcriptome of Cryptococcus neoformans var. grubii Reveals Complex RNA Expression and Microevolution Leading to Virulence Attenuation. *PLoS Genet.* **10,** (2014).
260. Chu, C. P. *et al.* RNA-seq of serial kidney biopsies obtained during progression of chronic kidney disease from dogs with X-linked hereditary nephropathy. *Sci. Rep.* **7,** 1–14 (2017).
261. Illumina. RNA-Seq Data Comparison with Gene Expression Microarrays. *White Pap. Seq.* 1–8 (2011).
262. Chen, Y., Toffaletti, D. L., Tenor, J. L. & Perfect, J. The Cryptococcus neoformans Transcriptome at the Site of Human. *MBio* **5,** 1–10 (2014).
263. Loftus, B., Fung, E., Fraser, J., Heitman, J. & Hyman, R. The genome of the basidiomycetous yeast and human pathogen Cryptococcus neoformans. *Science (80-.).* **307,** (2005).
264. Goebels, C. *et al.* Introns Regulate Gene Expression in Cryptococcus neoformans in a Pab2p Dependent Pathway. *PLoS Genet.* **9,** (2013).
265. Schurch, N. J. *et al.* Erratum: How many biological replicates are needed in an RNA-seq experiment and which differential expression tool should you use? *RNA* **22,** 1641–1641 (2016).
266. Liao, Y., Smyth, G. K. & Shi, W. FeatureCounts: An efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30,** 923–930 (2014).
267. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15,** 1–21 (2014).
268. Love, M. I., Anders, S., Kim, V. & Huber, W. RNA-Seq workflow: gene-level exploratory analysis and differential expression. *F1000Research* **4,** 1070 (2015).
269. Gene Ontology Consortium. The Gene Ontology Consortium. Gene ontology: tool for the unification of biology. *Nat. Genet.* **25,** 25–29 (2000).
270. Lomax, J. Tutorial section Get ready to GO ! A biologist ' s guide to the Gene Ontology. *Brief. Bioinform.* **6,** 298–304 (2005).

271. Alexa, A., Rahnenführer, J. & Lengauer, T. Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics* **22,** 1600–1607 (2006).
272. Supek, F., Bošnjak, M., Škunca, N. & Šmuc, T. Revigo summarizes and visualizes long lists of gene ontology terms. *PLoS One* **6,** (2011).
273. O'Meara, T. R. *et al.* The Cryptococcus neoformans Rim101 transcription factor directly regulates genes required for adaptation to the host. *Mol. Cell. Biol.* **34,** 673–84 (2014).
274. Clancy, C. J. *et al.* Cryptococcus neoformans var. grubii isolates recovered from persons with AIDS demonstrate a wide range of virulence during murine meningoencephalitis that correlates with the expression of certain virulence factors. *Microbiology* **152,** 2247–2255 (2006).
275. Pedroso, R. S., Lavrador, M. A., Ferreira, J. C., Candido, R. C. & Maffei, C. M. Cryptococcus neoformans var. grubii - Pathogenicity of environmental isolates correlated to virulence factors, susceptibility to fluconazole and molecular profile. *Mem. Inst. Oswaldo Cruz* **105,** 993–1000 (2010).
276. Monk, B. C. & Goffeau, A. Outwitting multidrug resistance to antifungals. *Science (80-.).* **321,** 367–369 (2008).
277. Sanguinetti, M. *et al.* Role of AFR1, an ABC transporter-encoding gene, in the in vivo response to fluconazole and virulence of Cryptococcus neoformans. *Infect. Immun.* **74,** 1352–1359 (2006).
278. Hu, G., Cheng, P. Y., Sham, A., Perfect, J. R. & Kronstad, J. W. Metabolic adaptation in Cryptococcus neoformans during early murine pulmonary infection. *Mol. Microbiol.* **69,** 1456–1475 (2008).
279. Elowitz, M. B., Levine, A. J., Siggia, E. D. & Swain, P. S. Stochastic Gene Expression in a Single Cell. **297,** 1183–1186 (2002).
280. Ferrareze, P. A. G. *et al.* Transcriptional Analysis Allows Genome Reannotation and Reveals that Cryptococcus gattii VGII Undergoes Nutrient Restriction during Infection. *Microorganisms* **5,** 49 (2017).
281. Kelliher, C. M., Leman, A. R., Sierra, C. S. & Haase, S. B. Investigating Conservation of the Cell-Cycle-Regulated Transcriptional Program in the

Fungal Pathogen, Cryptococcus neoformans. *PLoS Genet.* **12,** 1–23 (2016).

282. García-Rodas, R. *et al.* Capsule Growth in *Cryptococcus neoformans* Is Coordinated with Cell Cycle Progression. *MBio* **5,** 1–13 (2014).
283. Nogales, E. Structural Insights Into Microtubule Function. *Annu.Rev.Biophys.Struct* **30,** 277–302 (2001).
284. Ko, Y. J. *et al.* Remodeling of global transcription patterns of Cryptococcus neoformans genes mediated by the stress-activated HOG signaling pathways. *Eukaryot. Cell* **8,** 1197–1217 (2009).
285. Maier, E. J. *et al.* Model-driven mapping of transcriptional networks reveals the circuitry and dynamics of virulence regulation. *Genome Res.* **125,** 690–700 (2015).
286. Panepinto, J. *et al.* The DEAD-box RNA helicase Vad1 regulates multiple virulence-associated genes in Cryptococcus neoformans. *J. Clin. Invest.* **115,** 632–641 (2005).
287. Banks, I. R. *et al.* A Chitin Synthase and Its Regulator Protein Are Critical for Chitosan Production and Growth of the Fungal Pathogen Cryptococcus neoformans A Chitin Synthase and Its Regulator Protein Are Critical for Chitosan Production and Growth of the Fungal Pathogen C. *Eukaryot. Cell* **4,** 1902–1912 (2005).
288. Kelliher, C. M. & Haase, S. B. Connecting virulence pathways to cell-cycle progression in the fungal pathogen Cryptococcus neoformans. *Curr. Genet.* **63,** 803–811 (2017).
289. Pukkila-Worley, R. *et al.* Transcriptional Network of Multiple Capsule and Melanin Genes Governed by the *Cryptococcus neoformans* Cyclic AMP Cascade. *Eukaryot. Cell* **4,** 190–201 (2005).
290. Chang, Y. C., Penoyer, L. A. & Kwon-Chung, K. J. The second capsule gene of *Cryptococcus neoformans, CAP64*, is essential for virulence. *Infect. Immun.* **64,** 1977–83 (1996).
291. Liu, O. W., Kelly, M. J. S., Chow, E. D. & Madhani, H. D. Parallel β-helix proteins required for accurate capsule polysaccharide synthesis and virulence in the yeast Cryptococcus neoformans. *Eukaryot. Cell* **6,** 630–640 (2007).

292. Geddes, J. M. H. *et al.* Analysis of the protein kinase a-regulated proteome of Cryptococcus neoformans identifies a role for the ubiquitin-proteasome pathway in capsule formation. *MBio* **7,** 1–15 (2016).
293. Giles, S. S. *et al.* The Cryptococcus neoformans catalase gene family and its role in antioxidant defense. *Eukaryot. Cell* **5,** 1447–1459 (2006).
294. Pascon, R. C., Ganous, T. M., Kingsbury, J. M., Cox, G. M. & McCusker, J. H. Cryptococcus neoformans methionine synthase: expression analysis and requirement for virulence. *Microbiology* **150,** 3013–3023 (2004).
295. Mogensen, E. G. *et al.* Cryptococcus neoformans senses CO 2 through the carbonic anhydrase Can2 and the adenylyl cyclase Cac1. *Eukaryot.Cell.* **5,** 103–111 (2006).
296. Idnurm, A. & Heitman, J. Light controls growth and development via a conserved pathway in the fungal kingdom. *PLoS Biol.* **3,** 0615–0626 (2005).
297. Laurenson, I., Ross, J. & Milne, L. Microscopy and latex antigen negative cryptococcal meningitis. *J. Infect.* **36,** 329–331 (1998).
298. Kimura, M., Kaufman, F. & Hashimoto, S. Pulmonary cryptococcosis due to a capsule-deficient strain confused with metastatic lung cancer. *Mycopathologia* **140,** 65–68
299. Cox, G. M. *et al.* Superoxide Dismutase In uences the Virulence of. *Infect. Immun.* **71,** 173–180 (2003).
300. Xue, C. *et al.* Role of an expanded inositol transporter repertoire in Cryptococcus neoformans sexual reproduction and virulence. *MBio* **1,** 1–14 (2010).
301. Cheong, J. W. S. & Mccormack, J. Fluconazole resistance in cryptococcal disease: Emerging or intrinsic? *Med. Mycol.* **51,** 261–269 (2013).
302. de Oliveira Schneider, R. *et al.* Zap1 regulates zinc homeostasis and modulates virulence in Cryptococcus gattii. *PLoS One* **7,** (2012).
303. Watkins, R. A. & King, J. S. Nutritional Requirements and Their Importance for Virulence of Pathogenic Cryptococcus Species. 1–20 (2017). doi:10.3390/microorganisms5040065
304. Cottrell, T. R., Griffith, C. L., Liu, H., Nenninger, A. A. & Doering, T. L. The

pathogenic fungus Cryptococcus neoformans expresses two functional GDP-mannose transporters with distinct expression patterns and roles in capsule synthesis. *Eukaryot. Cell* **6,** 776–785 (2007).

305. Thanh, L. T. *et al.* Multilocus sequence typing of Cryptococcus neoformans var . grubii from Laos in a regional and global context. *Med. Mycol.* **0,** 1–9 (2018).
306. Bódi, Z. *et al.* Phenotypic heterogeneity promotes adaptive evolution. *PLoS Biol.* **15,** 1–26 (2017).
307. Wolf, J. M., Rivera, J. & Casadevall, A. Serum albumin disrupts Cryptococcus neoformans and Bacillus anthracis extracellular vesicles. *Cell. Microbiol.* **14,** 762–773 (2012).
308. Fan, Y. & Lin, X. Multiple Applications of a Transient CRISPR-Cas9 Coupled with Electroporation (TRACE) System in the Cryptococcus neoformans Species Complex. *Genet* **208,** (2018).